DOCUMENT RESUME

ED 067 963 48 FL 003 684

AUTHOR	Berman, Arlene; And Others
TITLE	A Collection of Studies. Research on Syntactic Typology Volume III of III Volumes. Final Substantive Report.
INSTITUTION	Language Research Foundation, Cambridge, Mass.
SPONS AGENCY	Institute of International Studies (DHEW/OE), Washington, D.C.
BUREAU NO	BR-0-7721
PUB DATE	[30 Aug 72]
CONTRACT	OEC-0-70-4986-823
NOTE	472p.
EDRS PRICE	MF-$0.65 HC-$16.45
DESCRIPTORS	Arabic; Bengali; Case (Grammar); Comparative Analysis; Contrastive Linguistics; French; *Grammar; Interference (Language Learning); *Language Research; *Linguistic Theory; Malayo Polynesian Languages; Pronouns; Second Language Learning; Serbocroatian; *Syntax

ABSTRACT

This final volume of a three-part report on various language research projects contains ten studies on a wide variety of language issues. The topics include "Some Aspects of Language Interference," "Contrastive Analysis of Case Marking in Polynesian Languages," "Notes on the Formation of Complex Sentences in Tunisian Arabic with Emphasis on the Relative Clause," "On Serbo-Croatian Complement Sentences," "Outline of Bengali Grammar," "Number in the Bengali Noun Phrase," "Pronominalization and Equi-Noun Phrase Deletion in Bengali," "Language-Particular Rules and Explanation in Syntax," "Evidence for Shadow Pronouns in French Relativization," and "Induced Errors." For the first two volumes of the report, see FL 003 682 and FL 003 683. (VM)

ED 067963

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

VOLUME III
OF III VOLUMES

A COLLECTION OF STUDIES

"Some Aspects of Language Interference"

Arlene Berman and Michael Szamosi

"Contrastive Analysis of
Case Marking in Polynesian Languages"

Sandra Chung

"Notes on the Formation of Complex Sentences
in Tunisian Arabic with Emphasis on the Relative Clause"

Colette G. Craig

"On Serbo-Croatian Complement Sentences"

Colette G. Craig

"Outline of Bengali Grammar"

Sheila Jasanoff

"Number in the Bengali NP"

Sheila Jasanoff

"Pronominalization and Equi-NP Deletion in Bengali"

Sheila Jasanoff

"Language-Particular Rules and Explanation in Syntax"

David M. Perlmutter

FL 003 684

FILMED FROM BEST AVAILABLE COPY

Evidence for Shadow Pronouns in French Relativization"

David M. Perlmutter

"Induced Errors"

Nancy Stenson

Some Aspects of Language Interference

Arlene Berman
and
Michael Szamosi

0. This paper is a report on a project undertaken at the Language Research Foundation under Contract No. OEC-0-70-4986(823) with the Office of Education, in the summer of 1971. The outline of the report is as follows: Section 1 deals with preliminary background material for the project. Section 2 provides the intended aims of the research. In Section 3, we present our original plan of operation, its subsequent modifications, and the reasons for them. Section 5 is devoted to a detailed presentation of our findings. In Section 6, we outline our conclusions and present a general discussion.

1. The present project was undertaken in loose collaboration with various ongoing programs at Language Research Foundation (LRF), dealing with questions of second language acquisition. In particular, our project was intended to complement the work of Burt and Kiparsky (1971) and Stenson (1972), which deal specifically with the problems of teaching English as a second language.

In recent years, two main approaches to this question have been developed: contrastive analysis and error analysis. As noted in Stenson (and references cited there), contrastive analysis concerns itself with examination of both the learner's native (base) language and the language being learned (target language). Most ambitiously, it aims at predicting potential trouble spots by comparing the grammars of the base and target languages and pinpointing areas of conflict and/or agreement. The assumption is that various constructions in the target language will be more or less difficult for the learner, depending on the nature of those constructions in his native language and that textbooks should be designed with the base language of their users in mind. On a remedial basis, i.e. in terms of persons who have at least some control of the target language, the aim of contrastive analysis is to explain errors in the target language with reference to the processes of the base language or their interference; and on the basis of this, to propose effective ways of explaining to the language learner why he made those errors. The assumption is that a given mistake may have different explanations depending on the native grammar of the person who made it.

Contrastive analysis has not been widely exploited for various reasons. One is the paucity of even partial grammars of many languages, and thus the lack of material on which to base a contrastive analysis. Another factor is that teaching English as a second language, in this country, has worked largely on a remedial basis. Thus, Kiparsky and Burt deal with the situation of teaching classes of non-homogeneous students (e.g. foreign students at American universities), a situation in which attention to the native language of each individual student is impractical, especially given the lack of teaching materials based on contrastive analysis.

Error analysis disregards any possible base language interference, and deals simply with the incorrect English used by the language learner. As such, it is exclusively remedial, and in its strongest form presupposes that the language learner is at least consistent in his use of English. It aims at discovering the underlying system in the learner's incorrect use of the target language and comparing this 'grammar' to the 'standard grammar' of the target language. Thus, to a certain extent, it merges the situation of persons for whom English is a second language with that of people who speak 'substandard' dialects of English. A weaker form of error analysis has been developed by Burt and Kiparsky, which aims at classifying errors in terms of the degree to which they effect comprehensibility. While their approach does not presuppose internal consistency in the incorrect English of the language learner, it does bypass explanations of errors based on interference from the base language, and rely solely on evidence internal to the language learner's use of the target language.

2. Our project is part of an attempt to compare and evaluate these approaches. Our primary aim was to determine the possibility of correlating the errors and weak areas of language learners with their base language. Evidence of this type is necessary for establishing the validity of contrastive analysis. On the assumption that there are some types of errors that all language learners will make, regardless of the base language or the target language, and other errors that only some students will make, we wished to determine the extent to which these two types of errors could be separated. Such information will help provide answers to the question of whether a language can be taught 'on its own', as it were, ignoring the question of base language interference. Is it

necessary, desirable, or feasible to supplement the teacher of non-homogeneous groups by providing him with the main areas of interference between specific base and target languages, along with explanations arrived at through contrastive studies?

In addition, we wished to find out what types of evidence could be brought to bear on evaluating the above approaches, how could such evidence be gathered and interpreted, and whether there are types of errors which neither theory can deal with. As it turned out, this latter issue became rather significant.

It was in search of preliminary answers to such questions that we set out on this project. It is to be emphasized that the project was highly exploratory in nature. Our main purpose was to determine the feasibility and advisability of more comprehensive studies along the lines presented above, and to delineate, in so far as possible, specific problem areas worth following up. Since little research has been done to date in this area, one of our concerns was methodology. We tried several different ways of collecting data. We encountered many unexpected difficulties and pitfalls which provided us with some insight into how to go about doing such work. Many of these pitfalls, or failings of methodology, were not apparent until we began a detailed analysis of the data. We try to point these out as we go along, and hope that our suggestions and descriptions of difficulties peculiar to India will prove helpful to future researchers.

3. India was chosen as a suitable place for research of this sort because it contains peoples with many different base languages (fourteen are recognized officially and there are many others). Further, it is one of the few countries with an extensive and well-established program of English teaching, and a social attitude that attaches some importance to being able to speak English.

Our original plan was to survey as many of the language groups as possible. We intended to include at least one of the northern languages (i.e. one of Hindi, Punjabi, Marathi, Gujarati, Bengali, etc., which, with the possible exception of Bengali, may be considered almost identical, at least in terms of syntax) and one or more of the Southern (Dravidian) languages (Tamil, Telugu, Kannada and Malayalam), which are more different among themselves than those of the northern group. Further, we had planned to survey a language of the Dardic group (Kashmiri) which, while Indo-European and thus genetically related to the northern languages, is sufficiently separated as to be wholly distinct today.

Our intention was to visit various schools in the different language areas, taping classes of varying grades of proficiency in English, and, if possible, collecting samples of written material from them.

4. In many ways, the plan outlined above proved impratical. The difficulties of getting about in India turned out to be much greater than we had supposed, particularly in the (summer) monsoon season. In addition, the upset political situation in Kashmir and West Bengal (at the beginning of the Indo-Pakistani tension) prevented us from going to either of those places.

The above difficulties, combined with personal health problems, resulted in our limiting the survey to two major areas: New Delhi in the north and Bangalore in the south.

In New Delhi, we were in contact with the British Council, who proved very helpful in obtaining permission for us to visit classes in several schools. All the schools we visited had Hindi as the medium of instruction, with the exception of one, where classes were conducted in English. In the Hindi-medium schools English is taught as a subject from the seventh standard (about the seventh grade). All of the children we saw in these schools had studied English for at least two years. (Some, of course, may have learned some English previously at home.) The English-medium school was visited for purposes of comparison. The native language of most students in New Delhi is Hindi, with a substantial amount of Punjabi and a sprinkling of the other nothern languages represented. All students were fluent in Hindi.

In the south, our choice of Bangalore was motivated by two main factors. First, it houses the Southern Regional Institute of English, an establishment maintained by the four southern states (Tamilnadu (Madras), Mysore, Kerala, and Andhra Pradesh) and advised by the British Council. The Regional Institute provides training and remedial courses for the English teachers of these states as well as preparing English course materials for the southern schools. The Institute provided us with the opportunity to interview many teachers, as well as helping us contact a number of schools in the area.

Second, Bangalore is centrally located with respect to the four southern states and has a population that is guite heterogeneous linguistically. While its official language is Kannada (that of Mysore State), it contains substantial numbers of Tamil and Telugu speakers and some native Malayal's. Thus, in addition to Kannada-medium schools, there are schools in which Tamil or Telugu is the primary medium of instruction.

As it turned out, however, this situation proved to be a mixed blessing. The schools were not segregated linguistically and owing to different social, economic, and demographic factors within the city, most schools ended up with several groups of different speakers. This meant, of course, that we weren't always able to separate our subjects according to their base languages. Students questioned as to their native language and the language they used most frequently often named two or more languages, since many people in Bangalore are multilingual, both because of the proximity of the different linguistic communities and as a result of the frequent intermarriages among these communities. Thus, we cannot in general refer to specific errors or mistakes made by e.g. Tamil as opposed to Kannada speakers. For our data from the south of India, we use the blanket term 'southern' when we find reason to refer to it as distinct from the 'northern' or New Delhi samples.

We also encountered difficulties with our original methodology. As stated above, we had planned to sit in on, and tape, English classes. There were several obstacles to this. The first problem, a technical one, was the excessive amount of noise present, which resulted in very low intelligibility of the tapes. Because of the climate, most Indian classrooms are built in a very open way (sometimes no more than a thatched roof and cane walls on the sides). Even in brick buildings, classes are conducted with big windows wide open. Hence, our early recordings are full of extraneous street and animal noises.

The second problem was more basic, and had to do with the way classes are conducted in India. There is a strong emphasis on rote learning and almost no use of free discussion or original writing. At the beginning of our work, we often went home with no more than a few drill exercises from the pupils on the tape. Even when we could convince the teacher to encourage discussion, we mostly got monosyllabic answers from the students.

Accordingly, we changed our method to a more direct one. Instead of taping the classroom situation, we selected a few students at a time, and in a close setting, tried to engage them in conversation. Beyond the initial problems of shyness and finding a suitable topic that would allow them to contribute significantly to the conversation (e.g. an account of local festivals), this worked out quite well.

We were usually able to obtain written compositions from a good number of students in every class, so we ended up with a quite respectable amount of conversational (taped) and written material.

5. In this section, we will discuss our findings, according to the following schema. First in (5.1), we will address ourselves to those few areas where we found systematic errors or deviations

from Standard English which correlated with a given base language.

Next, in section 5.2, we will present the types of errors, systematic and random, which occur across-the-board in all of our samples. For the most part, these errors do not seem to reflect base language interference. We give a sample of these and engage in a limited discussion of their significance.

Lastly, in section 5.3, we will try to call attention to errors and confusions of the kind that seem to be unclassifiable in terms of the two above-mentioned theories, and discuss in general problems of how to elicit meaningful data and interpret it.

5.1 The number of errors, or deviations, where we found significant correlations with the base language was exceedingly small, considering the size of our corpus. Some of the reasons for this will become apparent in our discussion of general problems with the data in section 5.3. In the meantime, we will give a representative sample of such errors and point out their possible sources.

Among the errors found exclusively in the northern samples were the following (spelling is normalized in all quoted samples):

I. A pervasive use of adjectives as substantives, as in

(1) He is a very good in English
(2) His nature is a good
(3) He is a(n) obedient
(4) The canal was a big and full of waters
(5) Prabha is a brilliant and stand(s) first in (her) class
(6) She is a very simple and always wear(s) very clean clothes

Some of these errors may be due to simple mistakes or omissions in printing. Even an American pupil of 10 or 12 may be expected occasionally to omit a word from a sentence, thus (1) above instead of He is a very good student in English. But the large number and consistency of these errors indicate that they are not simple omissions. It is not entirely clear what they should be attributed to. Similar errors occur in the speech of many other non-native speakers of English (cf. Yiddish-English: He is a learned = he is a learned man). By and large, only people whose native language permits this kind of substantivization of adjectives, or people in whose native language identity-of-sense pronominalization operates by deletion, will make this type of error. (These two features, substantivization and identity-of-sense deletion, seem frequently to go together. Contrastive studies might be helpful in delineating the extent to which these features are correlated.) Hindi, Bengali, and presumably the other northern languages, are apparently like this. There is, however, another possible explanation for errors like those in (1) - (6): none of the northern Indian languages have an article. Thus, the students who made these errors may simply have not mastered the intricacies of the English article, and having learned, e.g. he is a doctor for their base language construction he is doctor, they may have generalized this to placing an article in front of every non-verbal predicate.

II. Another phenomenon restricted to northern Indian English is the 'excessive' use of the English complementizer that in quoting direct or indirect discourse (i.e. its use in places where Standard English would not have it):

(7) Then I came to my mother and asked her that can I go to visit zoo

(8) I came to my father and asked that I want to go to visit the zoo. My father ordered me that you want to go zoo to go but first you ordered your mother

(9) My friend asked me that I go to my father (= My friend asked me if I had gone to my father)

(10) Firstly, I would like to tell you that, that I am an optimistic type of boy

(11) As you know, that I am a Bengali, and as far as I know that my favourite festival is Durga Puja.

Sentence (11) may not belong with this group, properly speaking, since it may be the result of overgeneralization, or of an incorrectly administered drill. If the student learned that the correct form of English was I know that S, he may have generalized, incorrectly, to cases like as you know and as far as I know. The origin for the extra that in the other cases is not hard to find. In most northern Indian languages, there exists a special particle, placed either before, after, or inside a sentence (depending on the language) which indicates that that sentence is a direct or indirect quote. It would seem that the students, upon learning the English complementizer that, which appears, among other places, after many verbs of saying, knowing and believing, equated it with their own discourse-particle, and generalized its use in English to all the uses of their 'equivalent' particle.

III. Another systematic error among northern students was the consistent use of to before the objects of verbs of 'liking':

(12) I love to him very much

(13) On Sunday, I often enjoy to a picnic

(14) I am like to him (= I like him)

(15) He loves to studies

This is interesting, because as far as we have been able to establish, Hindi (and at least Bengali) use such a case for the objects of 'psychological predicates' like like. (The actual expression is more like of me (it) likes to him.) Quite clearly, what we are dealing with is simple base language interference.

IV. In the southern samples, curiously enough, the errors that were self-consistent had to do with verbs of discourse again. We found several examples of the following types:

(16) The old woman told do you like to go to the ball? Yes said Cinderella. Then she told get the six mices and one pumpkin.

Apart from the 'wrong' use of say and tell, which are problematical for most learners of English regardless of their base language, the peculiarity here was the distribution. From the data, it seemed to us that, whenever a conversation was described, one participant was reported as saying and the other as telling. We have no explanation for this phenomenon. The hypothesis that suggests itself is that these speakers treat say and tell as suppletive variants whose semantic values differ only in 'orientation' (toward one speaker or the other) and whose syntactic properties are identical. It follows from this that tell in these examples is always without an (indirect) object. We don't know whether a say-tell pair as described above exists in the Dravidian languages, so we can't say anthing about base language interference. Whether or not it does, the type of problem where non-native speakers perform slight semantic shifts of the lexical items of the target language is well worth following up.

V. The other mistake which occurred frequently in the South but not in the North was the repetition of verbs of saying; they very often appeared in 'pairs' both before and after the reported discourse:

(17) They replied we have a servant said the stepsister

(18) Weeping bitterly, she said yes I love to dance very much replied Cinderella

(19) She told Cinderella come back within twelve o'clock said the woman

Again, we have no inkling as to what caused this error, which showed up rather frequently. It may be that a similar construction is used in southern Indian languages -- a repetition of a verb of saying, parallel to the northern discourse particle to indicate reported discourse. This usage appears to show up only where direct quotes are being cited.

In the first case discussed above, we found that contrastive analysis provided two potential explanations for the type of error illustrated by (1) - (6). If these errors do represent some type of substantivization process, it is likely that contrastive analysis would be most useful in accounting for its occurrence, and providing an explanation that could be used in explaining the source of the error to the student. If, on the other hand, it is simply a problem of overuse of the article, contrastive analysis has little to contribute beyond predicting, rather vacuously, that since the base language has no article, students are going to have trouble with English articles. In this case, error analysis can be relied upon to provide a better insight into the developmental processes of second language acquisition.

In the case of the second type of error discussed above, the misuse of the English that complementizer, it would, of course, be possible to arrive at a generalization concerning the errors illustrated in (7) - (11) without reference to the base language. But it seems to us that, with an awareness of the particular process in the base language, one is in a better position to find the correct generalization (since it seems to be molded on the base language) and to effect a correction, simply by pointing out the differences between the Hindi (or Bengali, etc.) discourse particle and the English complementizer. Quite often, in order to understand what something is, it helps to be told what it is not. Note that the error analyst might interpret the mistakes of (7) - (11) simply as lexical errors (ask for tell) plus a failure to distinguish direct from indirect discourse reporting. The true generalization about these errors seems to lie elsewhere.

In the third case also, error analysis as well as contrastive analysis can provide a generalization to account for these errors, but it is only contrastive analysis which will show that the generalization was not made randomly. It seems that, whenever such erroneous generalizations in the target language are lifted 'wholesale' from the base language, the contrastive analyst is in a better position to explain that portion of the learner's internalized grammar of the target language.

This example also points up the need for a systematic study -- along the lines of error analysis -- of the learners' English. Recall sentence (15) He loves to studies. One's immediate impulse is to take this to mean He loves to study. But an examination of the data would reveal that there are no similar instances, e.g. He tries (wants, etc.) to studies. What we are dealing with, of course, is a version of He loves (his) studies, with to erroneously added after love, as in all the instances of III above. Thus, if a teacher not familiar with this kind of systematic mistake makes the apparently obvious correction studies --- study, he may later be surprised to find sentences like He works hard at his study, a legitimate generalization of the correction the student received. This latter mistake would, in effect, be a 'teaching mistake', a result of insufficient acquaintance on the part of the teacher with the student's English.

5.2 In this section we present the kinds of mistakes which do not seem to correlate with any particular base language. They were found 'across the board', both in the southern and in the northern samples. In general, this does not rule out base language interference, since it may be that there are features common to both the northern and the Dravidian languages which caused some of these errors. (This is indeed the case in the phonological imperfections of Indian English, which we shall discuss later on.) There is good reason, however, to believe that many of these mistakes are not the result of interference. In most cases, the errors seem to be random, with the same speakers exhibiting both correct and incorrect forms of the same construction.

All of the errors dealt with here are what Burt and Kiparsky called "local", i.e. they do not involve the entire logical or semantic structure of the utterance, but are restricted to simple 'surface' phenomena such as the use of articles, prepositions, or correct morphological forms of given lexical items. In what follows, we shall list the most pervasive ones and present a discussion of them at the end of this section.

I. We have pointed out, in the above section, a particular misuse of the English indefinite article a; that it was used very often before adjectival predicates in our northern samples. Apart from this consistent misuse, it seems that Indian speakers are confused about English articles in general. In countless instances we found no article where one was needed (All of the examples in this section are chosen randomly from northern and southern samples, both oral and written, without note of source):

(19) She was beautiful girl
(20) His father is doctor
(21) He is best student in class
(22) The next cage was lion cage
(23) We saw bear
(24) I go with him in morning for morning walk

In other cases, we found the wrong article, or an article in the wrong place, or an article where none was needed;

(25) He is the player of football
(26) They are old(er) than the Cinderella
(27) Cinderella had a two stepsister
(28) They will not take food till the dusk
(29) The many people come to dance
(30) Prablha is a my best friend
(31) He is respect the all teachers
(32) He is the our class monitor

(33) He answered their all the questions

(34) This festival is a great fun

(35) She bring her a largest pumpkin

It is quite clear from the above representative sample that these students had absolutely no idea what the article in English was. These errors do not follow any systematic pattern; the same speaker or writer would commit any number of them, in addition to using the article correctly. Note, by the way, that examples such as these make it unlikely that the type of error discussed in 5.1 above represents an error in using the article, since if that were the case, we would expect to find an equal number of 'accidentally' correct examples, in which the article was absent, parallel to (19) - (22).

II. The use, or rather, the non-use, of plurals in nouns was also an overall problem. Plural forms were more often omitted than not, both with and without numerals:

(36) All teacher proud of him

(37) The wife had two daughter

(38) In the zoo, we saw the animal

At other times, the plural was used when it was not required:

(39) On that day, peoples celebrate this festival

(40) He gave him one rupees

In the following example both types of error are found:

(41) Six white mices turn into six white horse

Here again, the incorrect form was common both in the northern and southern samples, and there seemed to be no law governing the distribution of correct and incorrect forms.

III. Possessive 's also posed a problem. In the few instances where the context required it (in our samples), it was absent:

(42) He is my father friend
(43) Arun manners are very good
(44) His mother name is Kamla

Possessive pronouns (his, her, my, etc.) were properly used in the same environment, and we noticed that many children used the longer possessive construction with of (e.g. the mother of Cinderella) rather than the more common (in Standard English) one when a full noun was the possessor.

IV. The use of prepositions was likewise problematical. Above, we pointed to a systematicity in the use of to before objects of verbs like like in the northern samples. There were other subregularities indicating that at least some of the uses of prepositions were direct translations from the various base languages, but, again, their use seemed to be too erratic to point to a real systematicity. Among the examples we found were the following:

(45) The shoe fitted to her foot
(46) We came the lion house
(47) He comes to school at time
(48) My mother wait me
(49) Every day we go to morning walk
(50) He always help to the weak boys
(51) We return to home
(52) In the evening, we returned our home by bus
(53) We went with me the zoo

Note that many of these examples involve motion verbs of some sort, and the errors here may well be induced by the teaching method. If children were told that one says: we go to the store, we go home, they may easily have missed the generalization that the to form is normal, with home being exceptional. If, as seems most likely, objects of motion verbs are marked with to in their own languages, the lack of to with home may have struck them as especially noteworthy. This may have led to the generalization of the pattern without to. The use or nonuse of prepositions is, in most languages, including English, fairly idiosyncratic, and it is not surprising that children mix up forms. Neither contrastive analysis nor error analysis seem to offer much in terms of this sort of mistake.

V. A very pervasive problem, especially in the south, was an absolutely random mix-up of masculine and feminine pronouns. The following is a fair example:

> (54) Cinderella was a beautiful girl. But she was poor. He had two step sister. She had comb their hairs and dress them. He had do household work. The king of the country arranged a ball and invited many people to come and dance to marry her son with prettiest girl in land.

The above summarizes the most common errors found in the nominal construction. We shall defer a discussion of this until the end of this section and will proceed now with the most frequent errors in verbs and verbal phrases.

VI. By far the most common (and least serious) mistake was the lack of agreement (i.e. lack of -s on third person singular verbs). To a lesser extent, but still rather pervasive, past tense -ed was absent. Interestingly enough, agreement showed up in the form does, when this auxiliary is used. The reason for this is probably the same as that for the overuse of irregular forms which we discuss below.

VII. In strong verbs, we found that very often, even in the present tense, the past tense, or strong form was used, whereas, for the regular past tense, the tense marker, -ed, was often left off. Thus, we find, in the same passage:

> (55) He help me in my study. He is the best boy of our class. He stood first in eight class. He won many prizes.

It is clear, from the rest of the passage, that all the verbs should be in the present tense (except, possibly, the present perfective has won for won). Note that help, which is morphologically regular in all tenses, is correctly given in the present tense, while stood and won, which are morphologically irregular in the past, are given in the past. In general, it seemed to us that morphologically irregular forms, regardless of their function, were often given instead of the regular forms. This sort of error is very likely induced by the way in which English is taught. Grammar books normally devote a great deal of space to listing exceptions, as compared to stating generalizations. Thus, typically, a textbook will present the standard past tense formation in a sentence, and then devote a page or more to listing the exceptional forms. Students practice the exceptional forms through extensive drills and repetition and end up remembering them best and therefore using them everywhere. The fact that the forms of be, (am, is, are, was, etc.), which are highly irregular, are used more or less correctly by most students tends to support this claim.

VIII. A more 'serious' problem is the very pervasive use of _be_ as an auxiliary in simple tenses:

(56) On Sunday I am go to the LalBagh

(57) Cinderella was ran away (from) the ball

(58) My sister is say his mother aunt (= my sister calls his mother aunt)

(59) He is always speak the truth

(60) We are play hockey every day

Connected with this, we feel, is the use of the progressive _be_+_ing_ form for the simple present:

(61) He is living next door to my house

(62) He is having two daughters

(63) He is driving a scooter

In all the above, it is clear from the context that the proper form is the simple present. This participial form was pervasive in all our samples, and, in fact, is fairly standard in the speech of Indians who otherwise speak excellent English. We have been able to ascertain that the habitual present is expressed by a periphrastic construction involving an auxiliary and a tenseless form of the main verb in Hindi and Bengali (and presumably all the other northern languages). We do not know whether this is the case in the south.

The above list does not exhaust the kinds of errors found. There were numerous simple lexical errors and mixups, and some of the compositions were totally unintelligible. In all areas, children displayed a fairly wide range in competence. We have tried to give here a representative sample of the most frequent errors, the kind that were most characteristic in both the written and spoken English of Indian students, regardless of their base language.

Some of the mistakes can be directly attributed to the idiosyncrasies of English. Thus, it seems that mistakes in the correct form of the verb (VI) will be made by any learner of English, Indian or not, because there is a great deal of irregularity in the morphology of strong verbs. The particular form of the errors in English (i.e. won for win, rather than winned for won) seems to be due to excessive rote learning, so that the best remembered forms are the irregular ones, which are therefore used most often. But it is likely that this area of English will cause difficulties of one sort or another for almost all learners of English.

Another area in which the idiosyncrasies of English are likely to produce errors is the use of prepositions (IV), although here base language interference seems to play a role in determining the types of errors made. While mistakes of this kind were made by both northern and southern Indians, speakers of given languages were fairly consistent in their mistakes.

An error in which base language interference seems definitely to be involved is the use of be as an auxiliary (VIII above) in forms like I am go(ing) for I go, since, as noted, a periphrastic form for present habitual action is common to the Indian languages.

The other errors listed above don't seem to lend themselves to explanations along the lines of contrastive analysis. Nor does simple error analysis help here. It is possible to say that the frequent absense of the third person marker -s is the result of 'over-generalization' (a term in error analysis, which indicates that the student, having found no agreement markers on five out of the six personal forms of the verb, arrived at the natural generalization that there is none on the sixth). But it just doesn't ring true. The fact that, quite randomly, one does find verbs marked with -s, and without it, within the same sequence, and our experience in talking to these students and reading their compositions point to the simple conclusion that the reason for these mistakes is plain confusion. Rather than overgeneralization, or arriving at a wrong generalization on the basis of the data, it appears that the students made no generalization at all. In some areas, the facts of English have not succeeded in impressing themselves on the students' minds, resulting in the random use of the forms they have heard.

The same seems true of the erratic use of plural forms and of personal pronouns. In the latter case, base language interference may have played some part. The confusion about the gender of pronouns was more prevalent in the south. This is probably because some of the southern languages (at least Tamil) have no gender distinction. But, in general, most of the errors in this section are what one might call 'developmental errors', errors indicating that the relevant areas of English simply have not been even partially mastered by the students. There are also indications that some of these 'mistakes' are simply learned, or taught that way. This will be brought up in the next section.

5.3 In this section we will be concerned with mistakes that lie outside the scope of the traditional methods of analysis and with problems involving both the data and the methods of data collection.

We called attention, above, in connection with the taping of classes, to the lack of student involvement and classroom participation. The same problem showed up in many of the written samples we obtained. It appeared that, in many instances, the children had never been asked to write an original essay before (at least in English) and had no idea how to go about it. This was reflected in their compositions in various ways.

In one instance, a class of seventh graders was asked to write out the story of Cinderella. They had previously read and discussed the story. Apart from a few exceptions (one or two quite coherent accounts and one or two specimens of sheer word-salad), the batch we got was a very confusing one. In general, it contained a large number of the usual mistakes we listed above, but there were a few very curious ones. Taken out of context, or by looking at only one composition, they probably would defy explanation or be misinterpreted. It was only through comparing many of the compositions that we could arrive at a confident guess concerning their nature. They were errors due to imperfect memorizing of phrases from the original story, which occurred in constructions that the writer was wholly unfamiliar with, but which he thought he remembered well enough to reproduce. A good example of this is the following:

(66) Her stepmother uses tell many works to Cinderella and Cinderella has to do and she used to sleep near the fire

In this case, the problem is the interpretation of uses in the underlined phrase. Again, on the basis of other compositions, this seems to be an attempted rendering of something like

(67) The stepmother used Cinderella (as a servant), making her do the household work, etc.

It is notable that in the phrase she used to sleep near the fire the words used to are spelled youst, although in the underlined phrase in (66) uses is spelled correctly. Evidently, this latter uses was seen in the book and, although not quite understood, was remembered visually, but was not connected with the familiar and understood used to (youst) construction. Thus, what might easily be misinterpreted as a version of, e.g.

(68) Her stepmother used to give many jobs to Cinderella

turns out to be a mistake of a very different sort.

The above are examples of the kind of mistake that cannot be identified or stated systematically either as due to interference or as a result of a defective grammar underlying the speaker's English. They belong to an area that is outside the competence of both contrastive analysis and error analysis. Stenson points out that, quite often, classroom errors are unanalysable because they require a knowledge of what the student intended to say. Most researchers look at the data outside of the immediate situation in which the error was made, and are not in a position to identify its situational (rather than grammatical) source. Because we had a set of compositions based on a story which all had read, we were able to catch a glimpse of what was going on in the instances presented above. In general, the problem of differentiating between grammatical mistakes in attempts to use English more or less creatively and mistakes due to imperfect repetition of half-remembered phrases and constructs is a serious one for the linguistic analysis of second language errors, and researchers should be aware of this.

A similar case in New Delhi presented different problems. Here the students were given a choice of topics to write compositions about. Quite unexpectedly, an overwhelming majority chose to write about "My Best Friend". It turned out that practically everybody's best friend was Ram (if a boy) or Meena (if a girl), names which are equivalent to English Joe and Jane. Everyone's best friend's father was a doctor and mother a teacher. In fact, all the compositions on the topic (about sixty in all, from three separate schools) were virtually identical. Obviously, some such essay had appeared in one of the standard textbooks, and everyone was carefully reproducing it. These students had the story down pat and repeated, or attempted to repeat, it in short, graceless sentences that resemble those of first grade readers. The following is typical:

(69) A friend in need is a friend in deed. Ram is my best friend. He is a good student. He is my class fellow. He is of my age. His father is a doctor. He is the monutor of our class. He is good in every subject. He is a good player. He is very intelligent boy. He is captain of football team. He is ten years old. He is neat and clean.

Clearly, the creative use of English was at a minimum in compositions such as these, and they were rather useless as a source for 'interesting' mistakes.

Another difficulty that we encountered concerns the elicitation of meaningful, or interesting, data. As we pointed out, there were few 'global' errors in the above list of mistakes. These are errors that occur in complex sentences involving temporal conjunctions (before, while, etc.) or causal conjunctions (because, since, in spite of, etc.). We found few mistakes of this type for the simple reason that there were practically no such constructions in our examples. What had happened, in retrospect, is quite simple. The students were confronted with a situation where they had to impart information in a language which they had little command of. Further, many of them probably felt that they ought to 'put their best foot forward'. So they did what any reasonable person would have done, namely, they used the simplest sentences and constructions, the kind they felt they could handle with the least difficulty. Practically none of them risked getting tangled up in a complex because clause, or a conditional if-then clause. This complete restriction on the kinds of sentences used was the single most outstanding stylistic feature of all of our samples, both written and oral. In terms of subject matter, many of the samples were quite original, ranging from "My Hobby" to a description of marriage customs in the student's native village. But their use of the language showed little originality, and we ended up with an essentially uninteresting collection of sentences. We could not find out what kinds of mistakes Indians (northern or southern) make in relative clause constructions, in temporal conjunctions or the like, and our analysis revealed only the simplest, 'surface-level' mistakes. Other linguists may have had similar experiences and we believe that the problem will exist for future researchers. We are not quite sure what can be done about it. It is essentially a problem of methodology of collecting data. One approach may be to spend more time with small groups of individuals to relieve the nervous excitement of the situation, and even to steer the conversation onto subjects where the speaker is forced to do some complex explanation. This kind of thing is easier with adults than with children.

Lastly, we would like to touch upon the situation of English in India. We will be concerned only with those aspects that seemed to touch upon our study. English is spoken fluently, and in many cases, as a native language, by a large number of Indians, both in the north and in the south; among the local intelligentsia, the civil servants, and any number of well-educated, well-to-do families. The number of such people might be getting smaller, since it is now possible to receive a university education in India in at least six to eight of the Native languages -- one no longer has to be fluent in English to be admitted to a college. But most people above a certain socio-economic standing, who are middle-aged and above, and their families, speak English fluently, and a great many of them use it as their every day language. Their English (so-called Indian English) is regarded by most linguists as a distinct variety, or dialect of English, much like any other dialect (e.g. Scottish English). It has been studied to a certain degree, though not extensively, from the point of view of syntax. Its most obvious distinctions are phonological and lexical. What is relevant about it for our purposes is that, as far as the majority of Indians in India are concerned, Indian English is English. In one sense, this can be ignored. In a remedial teaching situation, as in remedial English classes at American colleges, the differences between Standard and Indian English are treated as mistakes or aberrations, and simply corrected -- the student is taught to use a different dialect of English much like speakers of Black English learning Standard English. In our study, we were able to ignore it, for the most part, because we were interested, basically, in the correlation of mistakes with particular base languages. In some instances, however, where we found across-the-board mistakes, our observations of Indian English have proven insightful. Thus, in 5.2 above, in section V, we discussed the use of be+verb for simple present tense (I am go = I go) in the samples we got from the students. We noted there that it may be due to the fact that such periphrastic constructions are common in Hindi and other Indian languages. Another, more probable, explanation lies in the fact that, as far as we have been able to observe, this usage is fairly common in Indian English, certainly among civil servants and school teachers, and not infrequent in the higher sociological strata. Thus, it may simply be that this misuse of be from the Standard English point of view is not a mistake on the part of the students at all, but part of the language the students are taught. In our interviews with the teachers receiving training at the Regional Institute of English, we found this usage very often. Obviously, this is not what the students' textbooks use, but they certainly hear it often enough from their teachers. Confirmation for this may be found in the fact that this misuse is pervasive both in the north and the south, although to our knowledge, only the northern languages use a similar construction.

This points up a general problem with analyzing factors of second language acquisition in a country where the teachers' native language is not the target language; it is likely to be a problem all over, but in India more so than other countries. If a better understanding of language teaching is desired, the problem certainly deserves further investigation.

6. In what follows, we would like to sum up and discuss some of the implications of our findings. As regards the objective of providing a clear-cut correlation between mistakes in the target language and the base language, our results are not very extensive and therefore not conclusive. Minimally, however, as a survey, our project does establish the existence of such correlations, and points to the definite desirability of further, detailed research along these lines. Such work, by researchers with a working knowledge of both the base and target languages, should be done on a more restricted basis, probably involving no more than two base languages, and perhaps only one. The question of the feasibility of introducing supplementary material based on contrastive studies of specific base and target languages into heterogeneous learning groups would be decided upon the outcome of such research. There is no doubt of its desirability and potential usefulness.

Our methdology, built from scratch, had many faults and stands in need of improvement. Yet much of what was wrong with it is, in fact, positive, as it is to be hoped that future researchers will take advantage of our failures and learn from our mistakes. The principal reason for not getting more extensive results seems to lie in the fact that, by and large, the data that we elicited was very restricted. As we pointed out, it lacked syntactic creativity, for the students, insecure in their command of English, fell back upon the simple constructions that they felt they could control. It seemed to us that much of the students' insecurity was a natural extension of the timidity that they show in the classrooms, based mainly on the fact that anything written in school is going to get graded, and that they were expected to perform, and perform well. It may be advisable, in the future, to try to work outside the educational system.

It would be desirable, further, to conduct long-term, follow-up surveys of the students' progress. It is to be expected that as the learner's competence developes, certain errors disappear and others come up in their place. The nature of these errors, and their sequence, could provide much useful information about second language acquisition, and would eventually enable teachers to foresee difficulties in their courses.

Another major obstacle appeared in the analysis of the data, which stemmed from the particular linguistic condition in India. Given that a student's mistake may represent his own imperfect learning, or that it may have been transmitted to him by a teacher, and that the teacher in turn may have been making a genuine mistake, or using Indian English, a significant analysis would require a long and careful identifying and peeling away of the several layers of errors. Obviously, on a survey like ours, this could not be done. The researcher has to familiarize himself with Indian English and with the English of his informants' teachers. If the above considerations are kept in mind, research can be carried out successfully.

As regards the issue of contrastive and error analysis, what becomes apparent from the study is that rather than being two competing approaches, these two approaches complement each other. We would like to devote some discussion to this, and illustrate it with an example or two.

The issues of foreign language teaching are of two kinds: those that relate to the practical problems of teaching, correcting errors, wrong generalizations and the like, and those that relate to theoretical questions and models of second language acquisition. Let us consider the practical side first, in situations where the learners have some competence in the target language already. It seems that here contrastive analysis has more to offer. If we review the kinds of errors found among non-native speakers of a language, we find that they can be divided into systematic and non-systematic errors. Most of the non-systematic errors are due to imperfect memory or performance slip-ups, or are situational mistakes that do not reflect on the student's competence. Others simply do not lend themselves to explanation. A large class of the non-systematic mistakes are lexical errors of a particular kind (e.g. tell for say, table for desk), where semantically related items seem to get mixed up. In a large number of these cases, it is possible to point to the source in the base language --

usually a lack of lexical differentiation in that language, e.g. not having two separate words for tell and say, or, if the two items exist, they are differentiated along semantic-syntactic parameters different from those operative in the target language. In these instances, contrastive analysis, as we observed in connection with several cases above, allows one to pinpoint the source of error and offers the possibility of meaningful correction, whereas error analysis can only point to the mistake but will not find a generalization which covers it.

With respect to systematic errors, the situation is similar. Given a systematic error, it is usually possible to establish the generalization underlying it. If the student's base language displays a similar generalization, this fact not only provides a source for the error, but, in addition, offers a way to affect the necessary corrections. This is especially the case when narrow idiosyncrasies of the base language come into conflict with idiosyncrasies of the target language.

A good example of this is the Indian English pronunciation of English /t d/. In standard English these stops are alveolar; in Indian English they are retroflexed. We can best understand the Indian English usage in considering the stop systems of the Indian languages: all of them possess two sets of coronal sounds, one dental and one retroflexed. On universal grounds, it seems that dental stops, rather than alveolar ones, are the most unmarked, or most natural, sounds. Hence, an Indian speaker, hearing an English alveolar /t/, will feel it to be divergent from the 'normal', dental sound, and therefore will reproduce it as equivalent to his own marked series, which is retroflex. A speaker of French, having only one /t/ series, will reproduce English /t/ as a dental. The retroflex stops of Indian English are so pervasive and so characteristic that one tends to think of them as an irremediable part of the way Indians speak. Yet it is entirely absent from the speech of an Indian when speaking French, since he can reproduce French dental /t/ perfectly well. In this case, an idiosyncratic fact about the Indian languages -- that they have two /t/ series -- and an idiosyncratic fact about English -- that /t/ is alveolar rather than dental -- result in Indian English sounding markedly different from Standard English, much more so than, say, Indian French from Standard French.

In a situation like this, any attempt to deal with the language learner's errors without regard to the native language of the learner is likely to be unsuccessful. A person dealing with Indian English without regard to the Indian languages would try to induce the learner to use alveolar /t/, a correction that would be extremely difficult for most learners. But given a knowledge that the Indian languages have a dental series, and noting that the substitution of dental /t/ for alveolar /t/ will produce a much more easily comprehensible version of English than the substitution of retroflex /t/ for alveolar /t/ would, contrastive analysis would suggest teaching speakers of Indian English to use dentals for standard English alveolars.

Note that this is a 'correction' that error analysis would be unlikely to arrive at, since nothing in the grammar of Indian English, where the use of retroflexed stops is quite consistent and where dentals do not appear, nor in the grammar of standard English, where dentals are nonexistent, would suggest such a correction. In instances like this, contrastive analysis obviously has the most to offer.

It is on the theoretical side that error analysis (along with contrastive studies) can be very useful. In the case of a systematic error, if no basis for it can be found in the base language, it becomes the task of error analysis to find out how and why erroneous generalizations were made on the basis of correct (but probably incomplete) data from the target language. It is the answers to such questions that provide empirical evidence for a theory of second language acquisition. In this sense, error analasis is analogous to the analysis of child language, which aims at generalizations underlying the process of language acquisition.

Contrastive analysis is also an important tool in this area. First, it can provide evidence which bears on general linguistic theory, by determining to what extent cross-language generalizations (interference) are made along the same parameters as the ones the linguist employs in constructing abstract grammars. The more correlations we find, the more we can be certain that linguistic theory is on the right track.

Second, contrastive analysis can aid in the construction of a theory of second language acquisition. It is not the case that every process, rule, or generalization in a given base language is going to interfere with processes in the target language. Establishing which language-particular areas of the grammar are most likely to cause cross-language interference provides important data for second language acquisition theories. Thus, both contrastive analysis and error analysis play a significant role in providing answers to important questions of foreign language teaching.

CONTRASTIVE ANALYSIS OF CASE MARKING IN POLYNESIAN LANGUAGES*

Sandra Chung

0.0 It has long been recognized that there are two basic types of case marking for a natural language, corresponding to the two usual ways of marking the actor and object of a transitive verb. The actor of a transitive verb can be identified with the subject of an intransitive; this accusative structure is demonstrated by Latin

(1)	puer-Ø puell-am amābat. boy nom girl acc loved	"The boy loved the girl."
(2)	puer-Ø fūgit. boy nom fled	"The boy fled."

Or, the subject of an intransitive can be identified with the object of a transitive verb. We find ergative structures like this in Niuean, a Polynesian language:

(3)	ne poka he tangata e akau. tns push erg man abs tree	"The man pushed (over) the tree."
(4)	ne fano e tangata. tns go abs man	"The man went."

Case marking is the most obvious rule which classifies noun phrases in this way, but it is not the only one. Any rule which refers to NP's may apply in an accusative or ergative pattern, and a language may be accusative in some respects, but ergative in others. In Latin, for instance, the object of a transitive nominalization resembles the subject of an intransitive one:

(5) neglegens religion-is "the neglecting of religion"
neglecting religion gen.

(6) adventus consul-is "the coming of the consul"
coming consul gen.

Compare this ergative pattern with the case marking of simple sentences in (1-2).

The fact that a language can have both types of case structure has led to some speculation about which type (if either) is more basic. Since the base component is presumed to be universal, it would be convenient if it prescribed the same NP classification for all languages, with patterns of the other type being derived at a more superficial level. Not surprisingly, most work in this area has suggested that the accusative is basic.[1] This is because even languages with ergative case marking make use of the notion of logical subject, and this notion is directly reflected by accusative structure.

In this paper I examine the claim that accusative is basic by considering case relations in several Polynesian languages. These languages are particularly appropriate for a contrastive study of case because they make use of both accusative and ergative case marking. Despite the fact that the ergative type is historically descended from the accusative, I show that neither type can be derived from the other on a synchronic level. Instead, both types are equally derived, and are related to larger systems of rules which are organized by case.

The outline of the paper is as follows: In §1, I describe four Polynesian languages, and show that their case marking forms a gradation from accusative to ergative. §2 examines the claim that all of these languages are underlyingly accusative, and that sentences in the ergative languages are derived by a passive rule.[2] The examination suggests that this is not so. Rather, the ergative is obtained historically through a reanalysis of the passive, and both types of case marking occur at the same level of derived structure. In §3-6, I show that my proposal is supported by the relation of case marking to other rules dealing with NP's. These rules are shown to differ systematically in the different languages, in a way which is determined or at least reflected by the case marking of simple sentences. Case then appears to be a principle which organizes rules, and the different types of case marking are representative of different rule systems.

1.0 Case Marking in Polynesian

Pawley (1966) divides the Polynesian languages into two major subgroups. These are Tongic, consisting only of Tongan and Niuean, and Nuclear Polynesian, which includes the other languages in the Polynesian Triangle as well as related languages in Micronesia (Outliers). Nuclear Polynesian is itself divided into a group of Eastern languages, and a less homogeneous group including the Outlier languages and Samoan. The four languages which we will consider are representative of these groups; they include Maori, an East Polynesian language, Pukapukan, an Outlier language, and Tongan and Samoan.

Each of the three subgroups exhibits a different type of case marking. This is remarkable, since Pawley's classification is based on morphology, and case marking in Polynesian is syntactic rather than morphological. Case is generally indicated by particles which precede the NP, much as tense or aspect is indicated by particles preceding the verb; e.g. Maori

(7) ka patu 0 te tangata i te poaka. "The man kills the pig."
tns kill nom the man acc the pig

In some Outlier languages, case can be indicated by position, with an unmarked actor always preceding an unmarked object; e.g. Sikiana

(8) t tamalaa ni kai te pakeo. "The man ate the shark."
the man tns eat the shark

These constructions operate differently in the different subgroups, with the Tongic languages being ergative, the Eastern languages accusative, and the Samoic-Outlier languages having a mixed type of case. Because case marking is distributed in this way, our four languages will represent different stages in the accusative-to-ergative continuum.

1.1 Maori. With few exceptions,[3] all sentences in Maori have some type of accusative case marking. In the active, the objects of transitive verbs are preceded by the particle i, while actors of transitives and subjects of intransitives are preceded by 0:

(9) i epa 0 te tangata i te kupenga.
tns throw nom the man acc the net
"The man threw the net."

(10) i oma Ø te tangata. "The man ran."
tns run nom the man

Actor and object both follow the verb, and may be permuted:

(11) i epa i te kupenga Ø te tangata.
tns throw acc the net nom the man
"The man threw the net."

The two word orders (9,11) are just as frequent in simple sentences, but VSO is preferred over VOS in more complex sentences (including indirect objects, etc.). We will see that this preference for subject-before-object is characteristic of all Polynesian languages.

Transitive sentences can optionally undergo the passive rule, which places a suffix on the verb and assigns different case particles to the NP's. The object is marked with the nominative Ø, and the actor with e:

(12) i epa -ina Ø te kupenga e te tangata. "The net was thrown by the man."
tns throw pass nom the net agt the man

As in the active, actor and object can be scrambled with no change in meaning, but there is a slight preference for the word order VOS (as in 12). The question of exactly how the passive affects word order will be taken up in §1.6.

While the passive has a semantic range like the passive in English (i.e. in indicating state/result/completed action), it differs from the English passive in two important respects: it occurs more often than the active, and applies to a larger range of constructions than in English. Intransitives, for instance, may undergo the passive if they are followed by a locative prepositional phrase:

(13) ka moe - a 0 te moenga e te tangata.
tns sleep pass nom the bed agt the man

"The bed is slept-on by the man."

Derived from:[4]

(14) ka moe 0 te tangata i te moenga.
tns sleep nom the man on the bed

And transitive imperatives must undergo the rule:

(15) 0 tua - ina 0 te raakau (e koe).
imp fell pass nom the tree agt you

"(You) fell the tree!"

Compare the active form of (15), which is ungrammatical, with examples of intransitive verbs in the imperative construction:

(16) *0 tua (0 koe) i te raakau.
imp fell nom you acc the tree

Intransitives:

(17) e tuu (0 koe). "(You) stand [up]!"
imp stand nom you

(18) 0 haere mai (0 koe) ki te whare.
imp go here nom you to the house

"(You) come to the house!"

Both of these facts (and especially the frequent occurrence of the passive) may have brought about the regularization of the passive suffix in Maori. The passive suffix in Polynesian is represented by -(C)(i)a, where the particular value of -C is determined idiosyncratically by the verb; compare Maori awhina-tia 'be helped' with hopu-kia 'be caught'. It has been shown[5] that -C originally belonged to the verb stem, but was reanalyzed as part of the suffix when Polynesian lost all consonants in word-final position. A further stage of this reanalysis appears in Maori, where one of the ten passive suffixes has become the unmarked form. -*Tia* can replace other, historically correct forms of the suffix, and is used to passivize derived verbs and borrowings from English.

1.2 *Tongan*. On the surface, case marking in Tongan is very different from case marking in Maori. Actors of transitive verbs are preceded by the particle *'e*. Objects of transitives, and subjects of intransitives were at one time marked by *'a*, but a recent account by Shumway (1971) states that this particle "has virtually disappeared in the spoken language."[6] The only place where *'a* survives seems to be before proper names. So we have:

(19) na'e ui *'e* he pulē *Ø* e fefinē. "The boss called the woman."
tns call erg the boss abs the woman

(20) na'e faangota *Ø* e tamasi'ī. "The child fished."
tns fish abs the child

where *Ø* is sometimes replaced by *'a* in written or formal speech.

Word order is less flexible than in Maori, since the actor must precede the goal if the latter is not a pronoun. (21-2) show that this order is obligatory:

(21) na'e ui *Ø* ia *'e* he pulē. "The boss called her."
tns call abs her erg the boss

(22) *na'e ui *Ø* e fefinē *'e* he pulē.
tns call abs the woman erg the boss

Finally, Tongan lacks a passive rule, as is normal for a language with ergative case marking. Notice, though, that the ergative *'e* resembles the agent of the passive in Maori. The resemblance is supported by the presence of a small class of verbs in Tongan, which are usually identified as 'intransitive', but take a Maori-like accusative structure. For example:

(23) 'oku 'ofa Ø e tangatā (*'i*) he ta'ahinē.
tns love abs the man to the girl

"The man loves the girl."

Like the other 'pseudo-accusatives', *'ofa* is a verb of emotion/perception, and looks like an intransitive followed by a locative phrase:

(24) 'e folau Ø e tangatā (*'i*) he vakā.
tns sail abs the man in the ship

"The man will sail in the ship."

Both constructions have recently eliminated the particle *'i* in casual speech.[7] As a result, the *'i* phrase is distinguished from an absolutive NP only by the form of the definite article (*e* for absolutives, *he* otherwise).

Sentences like (23) raise the possibility that Tongan is accusative in underlying structure, and that ordinary transitives like (19) are derived by means of a passive rule. Note that this rule must be obligatory for the vast majority of verbs, since sentences like (25) are ungrammatical:

(25) *na'e ui Ø e pulē (*'i*) he fefinē.
tns call abs the boss to the woman

Moreover, most transitives would have to take a -*Ø* passive suffix, since sentences like (26) are usually not allowed:

(26) *na'e ui-*a* *'e* he pulē Ø e fefinē.
tns call pass agt the boss abs the woman

Still, the possibility of this type of derivation is increased by the fact that many verbs like 'ofa have a derived transitive form. This transitive usually has a durational meaning, and consists of the verb stem plus a suffixed -(C)(i)a:

(27) 'oku 'ofe - ina 'e he tangatā Ø e ta'ahinē.
tns love trans erg the man abs the girl

"The man loves the girl
(all the time)."

(27) bears a striking and suspicious resemblance to the Maori passive (12).

However, not all 'pseudo-accusatives' have a transitive form, and it is unclear whether this derivation is a lexical or syntactic process. This question is discussed in more detail in §2.3.

1.3 Pukapukan. Pukapukan has two types of case marking, accusative and ergative, which are associated with different styles of speech. The accusative is thought to be correct, and is used in formal situations; it closely resembles the case marking of Maori. Consider:

(28) na patu Ø te taane i te wawine.
tns hit nom the man acc the woman

"The man hit the woman."

(29) na lele Ø te manu. "The bird flew."
tns fly nom the bird

The actor and object in (28) may be permuted, but there is a definite preference for VSO order. This order is apparently required when one (or both) NP('s) is a pronoun or proper name:

(30) na patu Ø i a Turi i te taane.
tns hit nom pers Turi acc the man

"Turi hit the man."

But:

(31) *na patu i te taane Ø i a Turi.
tns hit acc the man nom pers Turi

Passives are derived by attaching case markers to the NP's (e for the actor, Ø for the object), and adding a -(C)(i)a suffix to the verb. The idiosyncratic forms of the suffix are still used, but there are three variants which have been generalized: these are -ia, -ina, and -ngia, the last two of which are sometimes combined as -ngina. Once the passive has applied, the NP's are optionally scrambled. It is worth noting that word order is more flexible in the passive than in the active, since a passive sentence may undergo scrambling even when the NP's are pronouns or proper names. Moreover, there seems to be no preference for VOS or VSO order:

(32) na patu - a Ø te wawine e te taane.
tns hit pass nom the woman agt the man

"The man hit the woman."

(33) na patu - a e te taane Ø te wawine.
tns hit pass agt the man nom the woman

"The man hit the woman."

Compare the passive in Maori, where a VOS order is slightly preferred.

It is tempting to try to link this free (and therefore unmarked) word order to other facts about the passive in Pukapukan. The passive occurs at least as often as the accusative, and the two constructions do not differ significantly in meaning; both merely indicated a transitive action (and this is reflected in the translation). When pressed, my informants stated that the passive could imply completed action, but it is clear from texts and conversation that this is not generally the case. The two constructions are often interchanged in simple sentences, with (e) actors of passive sentences being gapped after (Ø) actors of accusative ones, and vice versa.

(34) mea ake Ø i a Te Malo i lua ipu, waapiki - a
do away nom pers T.M. acc two half put close pass

mai Ø naa pulu ki lunga o lua
here nom pl husk on top of two

ipu.. "Te Malo took the two halves,
half placed the husks around
the two halves.."

The ergative case marking is considerably more restricted. In this construction the actor is marked with e and the object with Ø, but the verb appears in its stem form:

(35) na patu e te taane Ø te wawine.
tns hit erg the man abs the woman

"The man hit the woman."

(36) na patu Ø te wawine e te taane.
tns hit abs the woman erg the man

The ergative clearly resembles the passive, and like it has a relatively free word order. Both VSO and VOS are allowed, as demonstrated by (35-6). It differs from the passive in being considered an element of casual speech. Although my informants used the ergative often in conversations, they were reluctant to admit it, and claimed that the ergative was not "good language" but "just the way people talk." This marginal status reflects the fact that the ergative has a limited syntactic scope. For instance, the passive rule may apply to intransitives which are followed by a locative prepositional phrase (cf. Maori):

(37) na lele Ø te taane i te paanga.
tns run nom the man on the mat

"The man ran on/over the mat."

(38) na lele - wia Ø te paanga e te taane.
tns run pass nom the mat agt the man

"The mat was run over by the man."

But ergative marking is ungrammatical for this type of sentence:

(39) *na lele e te taane Ø te paanga.
tns run erg the man abs the mat

The passive may also apply to transitive verbs of emotion, as in

(40) ko mina - ngia Ø te yua e ku.
tns want pass nom the water agt me

"The water is wanted by me."

derived from

(41) ko mina Ø au i te yua. "I want the water."
tns want nom I acc the water

But these constructions are ungrammatical in the ergative:

(42) *ko mina Ø te yua e ku.
tns want abs the water erg me

In addition, ergative sentences are not allowed to undergo some transformations which apply to passive or accusative sentences. For instance, Question Formation will not affect the actor of an ergative sentence:

(43) *ko ai na waka - tuu 0 te wale. (Erg.)
Q who tns caus stand abs the house

But the same rule is grammatical for the actor of a passive or accusative:

(44) ko ai na waka - tuu i te wale. (Acc.)
Q who tns caus stand acc the house
"Who built the house?"

(45) ko ai na waka - tuu - ina 0 te wale.(Pass.)
Q who tns caus stand pass nom the house
"Who built the house?"

Finally, the ergative is semantically marked with respect to the other two constructions, since it emphasizes the action of the verb. Thus (35) can be translated as "The man really hit the woman" or "The man hit the woman (once)." The passive and accusative, on the other hand, merely indicate that the verb is transitive, with no emphasis or added meaning.

These facts suggest that the ergative should not be derived by the passive rule, but should be introduced by a separate rule of case marking. This transformation places an e before the actor and a 0 before the object, and therefore resembles one part of the passive rule. This analysis is examined in more detail in §2.

1.4 Samoan. Samoan also has a mixed type of case, but this differs substantially from the case system of Pukapukan. As in Pukapukan, there are two unmarked constructions, but these are passive and ergative, not passive and accusative. Both passive and ergative use the case markers e (actor) and Ø (object); the passive is distinguished from the ergative by the presence of a -(C)(i)a suffix (or its generalized form, -ina) on the verb.

(46) 'ua fasi - ina e le teine Ø le tama.
tns hit pass agt the girl nom the boy

"The girl hit the boy."

(47) 'ua fasi e le teine Ø le tama.
tns hit erg the girl abs the boy

"The girl hit the boy."

(48) gives an example of an intransitive sentence:

(48) na sau Ø le tama. "The boy came."
tns come nom the boy

Both the passive and the ergative allow scrambling, and show a slight preference for the order VSO. However, my informants insist that VOS is just as grammatical:

(49) 'ua fasi - ina Ø le tama e le teine.
tns hit pass nom the boy agt the girl

"The girl hit the boy."

(50) 'ua fasi Ø le tama e le teine.
tns hit abs the boy erg the girl

"The girl hit the boy."

This same freedom seems to characterize the use of the ergative as opposed to the passive, since informants claim that the two are usually interchangeable. Semantically, both types indicate that the action of the verb is transitive; the ergative can emphasize immediate action, and the passive can emphasize completed action, but these distinctions are not usually observed. The constructions also have the same syntactic range; for instance, neither may apply to intransitive verbs followed by a locative phrase (cp. Maori, Pukapukan). While it is true that passives formed from intransitives are recognized as grammatical, they are dismissed as archaic and overly respectful speech:

(51) 'ua nofo Ø le faife'au 'i le nofoa.
tns sit nom the minister on the seat

"The minister sat on the seat."

(52) ?'ua nofo - ia Ø le nofoa e le faife'au.
tns sit pass nom the seat agt the minister

"The minister took his seat."

(53) gives the ergative version of (51):

(53) *'ua nofo e le faife'au Ø le nofoa.
tns sit erg the minister abs the seat

In general, it is difficult to find places where the passive is grammatical but the ergative is not, or vice versa.[8] Informants sometimes prefer the passive, but emphasize that this preference is stylistic (i.e., verbs are felt to be "fuller, more complete" with -ina than without).

The accusative is used less often than either of these types, because most verbs which occur in the passive construction also occur in the accusative. (In this sense, 'passive' is a misnomer, even though it is perfectly clear that sentences like (46) correspond to passives in other Polynesian languages.)[9] Instead, the accusative is limited to a small class of verbs of emotion or perception, as illustrated in (54):

(54) 'ua alofa Ø le tama 'i le teine.
TNS LOVE nom the boy acc the girl

"The boy loved the girl."

The accusative has a rigid word order which is in keeping with its limited use. My informants claim that VSO is used more frequently, and is much better Samoan, than VOS:

(55) ?'ua alofa 'i le teine Ø le tama.
tns love acc the girl nom the boy

"The boy loved the girl."

Accusative sentences may also be placed in the passive by adding the passive suffix to the verb, and changing the case of the NP's. Although the result is formally identical to (46), there is one important difference: the passive suffix may not be removed to form an ergative sentence like (47). In other words, accusative and ergative case marking are mutually exclusive:[10]

(56) 'ua alofa - ngia e le tama Ø le teine.
tns love pass agt the boy nom the girl

"The boy loved the girl."

But:

(57) *'ua alofa e le tama Ø le teine.
tns love erg the boy abs the girl

This distribution of types of case is what we might expect, and is what actually occurs in Tongan and (trivially) in Maori. In this respect, it is interesting that accusative and ergative are options for the majority of verbs in Pukapukan.

1.5 Accusative and Ergative. These descriptions lead to a typology of case marking in Polynesian which can be briefly summarized as follows. Two of the representative languages are basically accusative; of these, Maori has no ergative case marking, and Pukapukan uses the ergative in casual speech. The other two languages are basically ergative, with Samoan using the passive and ergative in free variation. Languages with a 'regular' ergative have an accusative which is restricted, and applies only to a closed class of verbs of emotion or perception. The languages thus form a continuum from accusative to ergative case marking:

	Maori	Pukapukan	Samoan	Tongan
Accusative	x	x	(x)	(x)
Passive	x	x	x	(x)
Ergative	-	(x)	x	x

Different types of case seem to be associated with different word orders, since order is free in accusative languages but relatively fixed otherwise:

	Maori	Pukapukan	Samoan	Tongan
Accusative	either	either[a)]	VSO	VSO
Passive	(VOS)	either	(VSO)[b)]	VSO[b)]
Ergative	-	either	(VSO)[b)]	VSO[b)]

---

a) VSO for pronouns or proper names
b) VOS if the object is a pronoun

( ) = stylistic preference

Notice that even those languages where VSO is required have a special environment which allows the opposite order. This suggests that free scrambling is the older feature in Polynesian, and that fixed word order is a more recent development.

The typology also leads to some observations about the case system of proto-Polynesian. It seems clear, first of all, that the ergative is not a part of this system. Although the ergative occurs in several Polynesian languages, it is relatively new in Pukapukan and is not found in the Eastern languages at all. Moreover, all ergative languages have some type of accusative construction, but (as Maori demonstrates) the opposite is not true. As the ergative can be derived from the passive on a morphological level, it is logical to assume that it is a late development in Polynesian. I will therefore assume that the ergative is historically descended from the passive. This does not mean, however, that it is synchronically derived by the passive rule.

The accusative can be reconstructed for proto-Polynesian, as it occurs in all of the languages that I have described. Since it is VSO in those languages where it is restricted, we can reconstruct T1 as the basic Accusative Rule:

(T1) Accusative Marking

$$\underbrace{\text{Tns} \quad \text{V}}_{1} \quad \underset{2}{\text{NP}} \quad \underset{3}{\langle \text{NP} \rangle} \longrightarrow 1 \quad \underline{\emptyset} + 2 \quad \langle \underline{i} + 3 \rangle$$

The fact that the accusative allows scrambling in languages where it is a regular construction can be explained by positing a scrambling rule:

(T2) Scrambling (optional)

$$\underbrace{\text{Tns V}}_{1}\quad \underbrace{\text{Case NP}}_{2}\quad \underbrace{\text{Case NP}}_{3} \longrightarrow 1\ \ 3\ \ 2$$

Scrambling is ordered after Accusative Marking, and applies to passive sentences as well.

1.5.1 Excursus on Word Order. As it is given above, Accusative Marking applies to strings which are VSO, since it marks the first NP after the verb as actor, and the second NP as object. This is clearly the dominant surface order of Polynesian, but there is some question as to whether it is the underlying order as well. It has been suggested (Chomsky, 1965) that the universal base includes the constituent VP, so that all VSO languages would be VOS in underlying structure. But this analysis does not seem to work for Polynesian. Let us examine some of the reasons why.

Pronominalization in most Polynesian languages[11] works from left to right. The anaphoric NP is replaced by a pronoun, which may be followed by an emphatic particle: this is wua in Pukapukan, lava in Samoan, anoo or ake in Maori. In general, there is no distinctive set of reflexive pronouns, and informants claim that there is no difference between reflexive and non-reflexive pronominalization. Consider this example from Maori:

(58) i whaka - riri Ø te pirihimana i a Hoone
tns caus angry nom the policeman acc pers Hoone

ki a ia anoo.
to pers him emp

"The policeman made John
angry at himself."
Or: "The policeman made John
angry at him."

In (58), the object of the ki phrase has been pronominalized, and may refer to the actor or (direct) object, since it is preceded by both. In (59), the direct object has been pronominalized instead:

(59) i whaka - riri 0 te pirihimana i a ia anoo
tns caus angry nom the policeman acc prs him emp

ki a Hoone.
to pers Hoone

"The policeman made himself
angry at John."
Not: *The policeman made him
angry at John.

Here ia anoo may refer to the actor (te pirihimana), but it no longer refers to Hoone. Because (58-9) differ only in the place of the pronoun, pronominalization in Maori must be left-to-right.

Pronominalization also works this way in Samoan and Pukapukan, as is demonstrated by the conjoined sentences below. In (60-1), the pronoun follows its antecedent:

(60) Sam. 'ou te mana'o 'i le teine, 'e te ita 'i ai.
I tns like acc the girl you tns angry rel

"I like the girl, but you
hate her."

(61) Puk. na waka - atawai 0 i a Turi i a Te Malo
tns caus thank nom pers Turi acc prs T.M.

laaua ma Yina, na waka - weuwi
they and Yina tns caus farewell

0 i a ana wua i a laaua.
nom prs he emp acc prs them

"Turi thanked Te Malo and
Yina, then he said fare-
well to them."

In (62-3), it does not:[12]

(62) Sam. 'ou te mana'o 'i ai, 'e te ita 'i
I tns like acc rel you tns hate acc

le teine.
the girl

"I like it, and you hate
the girl."

(63) Puk. na waka - atawai Ø i a ana wua i a
tns caus thank nom prs he emp acc pers

Te Malo laaua ma Yina, na
T.M. they and Yina tns

waka - weuwi Ø i a Turi i a laaua.
caus farewell nom prs T. acc prs them

"He (someone else) thanked
Te Malo and Yina, then
Turi said farewell to
them."

Again the question of whether two NP's are coreferential is determined by whether the antecedent comes first. As in Maori, pronominalization must proceed from left to right.

These facts become significant when we examine how pronominalization behaves in accusative sentences; for in all three languages it is the actor which pronominalizes the object, and not vice versa. For instance, (64-5) are reflexive sentences in Maori:

(64) i pupuhi Ø a Hoone i a ia anoo.
tns shoot nom prs Hoone acc prs him emp

"John shot himself."

Notice that (65) violates the Crossover Principle:

(65) i pupuhi i a ia anoo 0 a Hoone.
tns shoot acc prs him emp nom prs Hoone

"John shot himself."

But (66-7) are not reflexive:

(66) i pupuhi 0 a ia anoo i a Hoone.
tns shoot nom prs he emp acc prs Hoone

"He himself shot John."

(67) i pupuhi i a Hoone 0 a ia anoo.
tns shoot acc prs John nom prs he emp

"He himself shot John."

The Samoan data is complicated by a preference for VSO order, which is illustrated in the examples below. My informants claim that (68-9) are both coreferential, but that (69) is "not good Samoan:"

(68) saa va'ai 0 le 'aveloli 'iate ia lava i
tns see nom the truckdr. acc/prs him emp in

le faa'ata.
the mirror

"The truckdriver saw himself in the mirror."

(69) ?*saa va'ai 'iate ia lava 0 le 'aveloli i
tns see acc/prs him emp nom the truckdr. in

le faa'ata.
the mirror

"The truckdriver saw himself in the mirror."

(70-1) were accepted as "referring to two different people," although (71) was again "not as good:"

(70) saa va'ai 0 'oia lava 'i le 'aveloli i
tns see nom he emp acc the truckdr. in

le faa'ata.
the mirror

"He himself saw the truck-
driver in the mirror."

(71)?*saa va'ai 'i le 'aveloli Ø 'oia lava i
tns see acc the truckdr. nom he emp in

le faa'ata.
the mirror

"He himself saw the truck-
driver in the mirror."

This preference for VSO is even stronger in Pukapukan, where both VOS examples were rejected completely (v. §1.3). Only (72) was considered to be reflexive:

(72) na mimili Ø te wawine i a ana wua.
tns touch nom the woman acc prs her emp

"The woman touched herself."

(73) *na mimili i a ana wua Ø te wawine.
tns touch acc prs her emp nom the woman

Compare (72) with (74), where the actor is pronominalized:

(74) na mimili Ø i a ana wua i te wawine.
tns touch nom pers he emp acc the woman

"He himself touched the
woman."

(75) *na mimili i te wawine Ø i a ana wua.
tns touch acc the woman nom pers he emp

Note that coreference is not affected by surface order in languages which allow scrambling (e.g. Maori). Actors may pronominalize objects which precede them, as in (65), but objects may never pronominalize actors. Since pronominalization works from left to right, it cannot apply at the surface level. Instead, it must apply to a VSO structure, and the NP's must be rearranged later by a scrambling rule.

Now let us suppose that all Polynesian languages have a verb phrase in underlying structure; in other words, that the basic word order is VOS, and the VSO shape of accusative sentences is derived by an optional transformation:

(T0) Object Shift (optional)

$$\underbrace{[_{\mathrm{VP}}\ \mathrm{Tns}\ \ \mathrm{V}}_{1}\ \ \underset{2}{\mathrm{NP}}\ \ \underbrace{{}_{\mathrm{VP}}]\ \ \mathrm{NP}}_{3}\ \longrightarrow\ 1\ \ 3\ \ 2$$

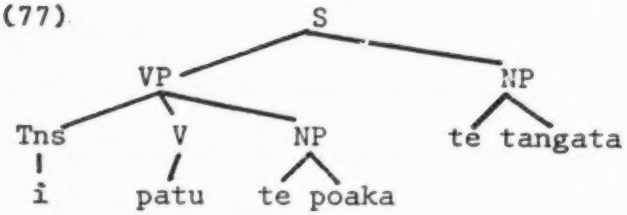
Object Shift will apply more often than not in languages where VSO is preferred (e.g. Samoan); it will be obligatory in languages where VSO is required (e.g. Tongan). It will also be obligatory when actor and object are coreferential, since sentences must be in the form VSO to undergo pronominalization. Finally, in languages which have both word orders, Object Shift will be distinct from the Scrambling rule, even though both do the same job of reordering actor and object. This is because Object Shift must precede pronominalization, but Scrambling must follow it (64-5).

In general, these properties of Object Shift argue that it is an unnatural rule. The situation is particularly clear in Maori, where the presence of Object Shift and Scrambling would mean that sentences could have ambiguous derivations. For example, (76) could be derived from a deep structure like (77) in two ways: either by applying Object Shift and Scrambling, or else by applying neither.

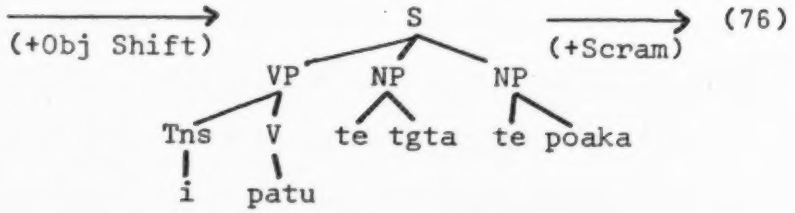
(76) i patu i te poaka Ø te tangata.
tns hit acc the pig nom the man

"The man hit the pig."

(77)



Deriv: (a)



(b) $\xrightarrow[\text{(-Obj Shift)}]{}$ $\xrightarrow[\text{(-Scram)}]{}$ (76)

All sentences with VOS order would be similarly ambiguous.

It thus appears that Polynesian languages do not have a constituent VP; rather, they are VSO in underlying structure, and this is reflected by the direction of pronominalization. This conclusion is supported by a number of other facts, of which I will mention only one. In Pukapukan, case marking can sometimes be omitted entirely, if the verb is in its stem form and the tense marker has been deleted. We would expect sentences in this minimal construction to reflect the word order of deep structure, since there are no case markers to indicate the underlying grammatical relations. In fact, these sentences are always interpreted as VSO. So maatou in (78) is always the actor, and never the object, of welaavei:

(78) welaavei maatou o maatou maatutua.
visit we poss we parents

"We visited our parents."
Not: Our parents visited us.

1.6 The Passive Rule. A passive can also be reconstructed for proto-Polynesian, since it occurs in all four languages which I have described. Moreover, it is marked with respect to the accusative in languages where both constructions are 'regular'. The passive applies a state or result in Maori; in Pukapukan, it can indicate completed action. The fact that it loses this special meaning in languages where the accusative is restricted suggests that its derived status is old. It is therefore this derived type which is attributed to proto-Polynesian. Since a deep structure for transitive sentences has already been reconstructed, we can obtain the passive by means of a transformation:

(T3) Passive Rule (optional)[13]

$$\begin{array}{llllcl} & \text{Tns} & \text{V} & \text{NP} & \text{NP} & \longrightarrow 1\ \ 2 + \underline{\text{(C)(i)a}}\ \ \underline{\text{e}} + 3\ \ \underline{\emptyset} + 4 \\ & 1 & 2 & 3 & 4 & \\ \text{Or:} & \text{Tns} & \text{V} & \text{NP} & \text{NP} & \longrightarrow 1\ \ 2 + \underline{\text{(C)(i)a}}\ \ \underline{\emptyset} + 4\ \ \underline{\text{e}} + 3 \\ & 1 & 2 & 3 & 4 & \end{array}$$

How the passive affects word order is a more complicated question. We have seen that languages with a regular passive allow it to undergo scrambling (e.g. Maori, Pukapukan, Samoan), but that these languages show different stylistic preferences for which NP comes first. For instance, Maori favors a VOS order for the passive, but Samoan favors VSO. As these stylistic preferences are not particularly conclusive, we must look elsewhere to determine how the passive actually works. Let us return to pronominalization in Pukapukan and Maori.[14]

Pronominalization is fairly complicated for passive sentences in Maori. Objects may be pronominalized by actors, exactly as in the accusative, and this operation is not affected by surface order:

(79) ka heru - ina e Hera Ø ia anoo.
tns comb pass agt Hera nom her emp

"*Herself will be combed by Hera."

(80) ka heru - ina Ø ia anoo e Hera.
tns comb pass nom her emp agt Hera

"*Herself will be combed by Hera."

Or, actors may be pronominalized by objects which precede them in surface structure. These constructions have the surface order VOS, but never VSO:

(81) ka heru - ina Ø a Hera e ia anoo.
tns comb pass nom prs Hera agt her emp

"Hera will be combed by herself."

Compare (81), which is coreferential, with (82), which is not:

(82) ka heru - ina e ia anoo Ø a Hera.
tns comb pass agt him emp nom prs Hera

"Hera will be combed by him himself."

The way that pronominalization works in these cases is not terribly straightforward, but seems to involve the two notions of primacy and precedence. NP's which are pronominalized by primacy (79-80) may undergo scrambling; NP's which are pronominalized by precedence (81-2) may not.[15] However, regardless of the details of this process, it is clear that some passive sentences (e.g. 81) must be VOS when pronominalization applies. (Recall that pronominalization works from left to right.) This means that the passive must reorder actor and object, since all transitive sentences are VSO in underlying structure:

(T3) Passive Rule (optional)

$$\underset{1}{\text{Tns}}\ \ \underset{2}{\text{V}}\ \ \underset{3}{\text{NP}}\ \ \underset{4}{\text{NP}} \longrightarrow 1\ \ 2 + \underline{(\text{C})(\text{i})\text{a}}\ \ \underline{\text{Ø}} + 4\ \ \underline{\text{e}} + 3$$

We find a similar situation in Pukapukan, where coreferential passive sentences may be VSO or VOS.[16] The only restriction is that pronouns must follow their antecedents in surface structure:

(83) na patu - a Ø te taane e na wua.
tns hit pass nom the man agt him emp

"The man was hit by himself."

(84) na patu - a e te taane Ø i a ana wua.
tns hit pass agt the man nom pers he emp

"*Himself was hit by the man."

In (83-4) the pronouns follow their antecedents. In (85-6), they do not:

(85) na patu - a e na wua Ø te taane.
tns hit pass agt him emp nom the man

"The man was hit by him himself (another man)."

(86) na patu - a Ø i a ana wua e te taane.
tns hit pass nom pers him emp agt the man

"He himself was hit by the (other) man."

Since pronominalization in Pukapukan works from left to right,

(83) suggests that the passive does reorder actor and object. It is possible that (84) (where the actor precedes the object) is derived by primacy, although this seems unlikely in view of the fact that (72) is not allowed.[17]

Both of these languages have regular accusative and passive constructions, so we would expect them to reflect the older features of the passive rule. As a result, the fact that the passive reorders actor and object can be reconstructed for proto-Polynesian. It is interesting that this derived status seems to break down in languages where the accusative has been restricted in some way. In Samoan, for instance, the passive may still reorder actor and object, but its preferred word order is VSO.

To conclude, I have reconstructed a deep structure for transitive sentences in proto-Polynesian. These sentences are subject to Accusative Marking, if they do not undergo the Passive Rule.

2.0 The Accusative-to-Ergative Drift

In §1, I claimed that ergative case marking is historically derived from the passive, so that (87-8) come from the same source in proto-Polynesian:

(87) Mao. ka kimi - hia Ø te tamaiti e Pita.
tns search pass nom the child agt Pita
"The child will be searched for by Pita."

(88) Ton. 'e kumi 'e Pita Ø e tamasi'ī.
tns search erg Pita abs the child

"Pita will search for the child."

One of the more straightforward explanations for this relationship is based on the fact that (87) has a passive suffix, but (88) does not. (The two examples also have different word orders, but this can be dismissed as a product of the Scrambling rule.[18]) If we assume that the suffix is deleted by a low-level transformation, then both types of sentences can be synchronically derived by means of the passive rule.

This argument has been advanced by Hale (1968) and Hohepa (1969), who claim that the Polynesian languages are gravitating towards some type of ergative case marking. Hale and Hohepa (hereafter H/H) state that the ergative is just an extension of the process by which passives are preferred in basically accusative languages (e.g. Maori). In ergative languages, that is, the passive is required for most transitive verbs; these then undergo a rule which deletes the passive suffix.[19] The H/H proposal is stated in speculative terms, but it is clearly intended to have one specific consequence. As Hohepa says, "ergative structures are not part of deep structure relations. Instead, they are produced by later syntactic processes, with passive [suffix] deletion following passivisation."[20]

The H/H proposal has several apparent advantages, of which the most obvious is that it explains the accusative-to-ergative drift. Although the accusative can be reconstructed for proto-Polynesian, the data of §1 suggests that it is being replaced by other constructions. For instance, the accusative is used less often than the passive in Maori; it is recognized as formal, not casual, speech in Pukapukan. (Recall that casual speech is the domain of the ergative in Pukapukan, a fact which suggests that the ergative is relatively new.) Moreover, the early texts and chants from several languages show that the accusative was used more often (and the ergative less) than it is now. In Pukapukan chants (c. 1930), the accusative is the normal type of case marking, and the ergative appears hardly at all. In Samoan texts (c. 1880), the passive occurs more often than the ergative, even though the two constructions are now used interchangeably.

The H/H proposal is also attractive from the viewpoint of linguistic theory, since it suggests how the ergative is connected to underlying grammatical relations. If the passive and the ergative are essentially the same construction, then the ergative is derived (and furthermore, is produced by a well-known transformation). This conclusion supports the view (advanced by Chomsky, 1965) that the relations of subject and object are defined universally. It also suggests how a language may change its case marking through time, since a shift from accusative to ergative would involve no more than a shift in conditions on the passive rule.[21]

However, the proposal is too general to apply to most

Polynesian languages in the form that it is given by H/H. In Tongan, for example, the passive is optional for verbs of emotion or perception, but must be obligatory for all other transitive verbs. The -*(C)(i)a* deletion rule is obligatory for most transitive verbs, but cannot apply to constructions where the passive is optional.[22] These facts are not a serious challenge to the H/H proposal, but they do suggest that it is too strong to account for the facts of Polynesian. In the following sections, I try to show some more compelling reasons why the proposal must be rejected for a weaker hypothesis.

2.1 *Preliminaries*. Let us suppose for a moment that the H/H proposal is correct, and that ergative case marking arose through an extension of the passive rule. This means that all Polynesian languages have the Passive in their list of transformations, and some of them have a rule of -*(C)(i)a* Deletion as well:

(T3) Passive Rule

Tns V NP NP → 1 2 + *(C)(i)a* *0* + 4 *e* + 3
1 2 3 4

(T4) -*(C)(i)a* Deletion

-(C)(i)a → 0

A rule of Accusative Marking will also be required, as most ergative languages use this construction for a limited class of verbs:

(T1) Accusative Marking

$$\begin{array}{cccc} \text{Tns} & \text{V} & \text{NP} & \langle \text{NP} \rangle \\ 1 & 2 & 3 & 4 \end{array} \longrightarrow 1 \quad 2 \quad \underline{\emptyset} + 3 \ \langle \underline{i} + 4 \rangle$$

The exact conditions on these rules will have to be specified for each language. Following Hale (1968), I assume that they take the form of lexical features, or redundancy rules.

It is worth remembering that the Passive has the same structure in all Polynesian languages, and varies only in the degree to which it is optional or obligatory. Moreover, this variation does not affect any other rules in the system, since the heart of the H/H proposal is that it limits the change to the passive rule.

If we follow the consequences of the proposal, we can make several predictions about languages in which the passive is required (or, almost required). Let us examine how these compare with the actual facts of Polynesian.

2.2 Word Order. First, we would expect ergative languages to have the surface order VOS, since this is the order produced by the passive rule. But the relationship between word order and type of case in these languages is actually quite different. As we have seen (§1.5), word order is flexible in languages with an optional passive, but it is more rigidly VSO in languages where the passive is required (e.g. Samoan, Tongan). This pattern is borne out by data from Niuean, which according to Hohepa (1969) has only the ergative construction. In Niuean, simple transitive sentences must be VSO. Sentences which seem to be VOS are interpreted as intransitives followed by a possessive phrase:

(89) ne poka he tangata e tama taane.[23]
tns push erg man abs child male

"The man pushed [his] son."

(90) na poka e tama taane he tangata.
tns push abs child male poss man

"The man's son pushed."

2.3 *Passive Suffix*. We would also expect the passive suffix -*(C)(i)a* to behave in one of two well-defined ways. In languages which are completely ergative, -*(C)(i)a* should be eliminated in a uniform way by the Deletion rule. Elsewhere the suffix should merely indicate that the verb is (underlyingly) transitive, since this is after all its function in accusative languages.[24] There are some languages, however, where neither of these requirements is met, because the syntactic function of -*(C)(i)a* has been reanalyzed as a semantic one. One language which has undergone this reanalysis is Tongan.

A few transitive verbs in Tongan need not undergo the Deletion rule. These verbs have different meanings, depending on whether they exhibit a -*(C)(i)a* in surface structure:

(91) na'e liaki 'e he tamasi'ī Ø e lesonī.
tns abandon erg the child abs the lesson

"The child threw away the lesson."

(92) na'e lieki - na 'e he tangatā Ø e fefinē.
tns abandon pass agt the man abs the woman

"The man deserted the woman."

These differences could conceivably be accounted for by a feature in the lexicon. But differences in syntactic properties cannot be explained so easily. For instance, *tanu* 'bury' is a transitive verb:

(93) na'e tanu 'e he tangatā Ø e kapā.
tns bury erg the man abs the can

"The man buried the can."

But its passive *tanu-mia* 'be completely covered' is intransitive:

(94) na'e tanu - mia (*'e he tangatā) Ø e kapā.
tns bury pass agt the man abs the can

Since the passive is supposed to produce a syntactically regular (and transitive) class, it cannot have been responsible for the -*(C)(i)a* in (94).[25]

The idea that -*(C)(i)a* is introduced by a different, lexical rule is confirmed by sentences where it appears, but cannot have been derived by the passive. For instance, *fehi'a* 'dislike' occurs in the accusative regardless of whether it takes the passive suffix:

(95) 'oku ke fehi'a ('i) he me'akaī.[26]
tns you dislike acc the food

"Do you dislike the food?"

(96) 'oku ke fehi'a - 'ia ('i) he me'akaī.
tns you dislike pass acc the food

"Do you abhor the food?"

(97) shows that the passive and ergative are not allowed:

(97) *'oku ke fehi'a - ('ia) Ø e me'akaī.
tns you dislike pass abs the food

Moreover, adjectives which take the passive suffix are often found in the accusative. This construction cannot be produced by a syntactic passive rule:

(98) 'oku ou maalie - 'ia ('i) he faivā.
tns I pleasing pass acc the movie

"I find the movie pleasing."

Derived from:

(99) 'oku maalie Ø e faivā. "The movie is
tns pleasing abs the movie pleasing."

(95-9) provide especially good evidence that -(C)(i)a is no longer introduced solely by the passive rule. Instead, it has become a meaning-bearing suffix which is probably derived by a lexical process. This shift in derivation is interesting in itself, but cannot be discussed in any detail here.[27] For our purposes, the most significant fact is that the shift is not predicted by the H/H proposal.

2.4 *Rule Interaction*. The H/H proposal does predict that the passive will not interact with other syntactic rules. This prediction turns out to be false, as we can see from two rather involved examples.

2.4.1 *Raising*. Sentences in Polynesian are subject to a number of transformations which are ordered after the passive rule. Most of these rules apply on a higher clause in the complex sentence, or are relatively superficial. In accusative languages, the rules usually refer to grammatical subject; that is, to the actor of an accusative verb, but the object of a verb in the passive. Consider the following examples of subject raising in Maori:

(100) kaahore i haere 0 te tangata. "The man didn't
neg tns go nom the man go."

When Raising has applied:

(101) kaahore 0 te tangata i haere. "The man didn't
neg nom the man tns go go."

Raising turns the subject of the lower sentence into the subject of the negative verb, kaahore.[28] (103) shows that the rule applies to the actors of accusative verbs:

(102) kaahore i patu 0 te tangata i te poaka.
neg tns hit nom the man acc the pig

"The man didn't hit the
pig."

When Raising has applied:

(103) kaahore 0 te tangata i patu i te poaka.
neg nom the man tns hit acc the pig

"The man didn't hit the
pig."

(104) shows that it does not apply to accusative objects:

(104) *kaahore (i) te poaka i patu (ai) 0 te tangata.
neg acc the pig tns hit rel nom the man

Raising also applies to passive sentences, but affects the object of the passive rather than the actor:

(105) kaahore i patu - a 0 te poaka e te tangata.
neg tns hit pass nom the pig agt the man

"The pig wasn't hit by
the man."

(106) kaahore Ø te poaka i patu - a e te tangata.
neg nom the pig tns hit pass agt the man

"The pig wasn't killed
by the man."

But:

(107) *kaahore (e) te tangata i patua(ai) Ø te poaka.
neg agt the man tns hit pas rel nom the pig

There are two obvious definitions for grammatical subject in Polynesian: either 'NP preceded by Ø', or 'NP closest to the verb'. If we assume that the first is correct, then we have:

(T5a) Raising I (optional)

$$\underset{1}{V}\ \underbrace{[_S\ \mathrm{Tns}\ \ V\ \ X}_{2}\ \ \underbrace{\emptyset\ \mathrm{NP}}_{3}\ \ \underbrace{Y\ _S]}_{4} \longrightarrow 1 + 3\ \ 2\ \ 4$$

If the second is correct, then the rule looks like:

(T5b) Raising II (optional)

$$\underset{1}{V}\ \underbrace{[_S\ \mathrm{Tns}\ \ V}_{2}\ \ \underbrace{\mathrm{Case}\ \mathrm{NP}}_{3}\ \ \underbrace{X\ _S]}_{4} \longrightarrow 1 + 3\ \ 2\ \ 4$$

Now H/H claim that the accusative-to-ergative drift consists merely of making the passive obligatory, and adding a -(C)(i)a deletion rule. If this is true, the SD's of other rules should not be affected by the change, but should keep referring to grammatical subject regardless of the conditions on the passive. This assumption has the interesting consequence that rules like Raising should apply ergatively in ergative languages. This is because objects of passive verbs will always satisfy the definition of grammatical subject ('NP closest to the verb' or 'NP preceded by Ø'), but actors of passive verbs will not (cf. T3).

However, the situation in ergative languages is actually quite different, since many rules like Raising have been reanalyzed to refer to logical subject. In Tongan, for instance, Raising still applies to the subject of an intransitive:

(108) 'oku lava ke huu Ø 'a Mele ki hono falē.
tns pos. tns enter abs prs M. to his house
"It is possible for Mary to enter his house."

Plus Raising:[29]

(109) 'oku lava Ø 'a Mele 'o huu ki hono falē.
tns pos. nom prs M. compl enter to his house
"Mary can enter his house."

But it affects the actor of a transitive (= passive) verb rather than the object:

(110) 'oku lava ke taa'i 'e Siale Ø e fefinē.
tns pos. tns hit erg Siale abs the woman
"It is possible for Siale to hit the woman."

(111) 'oku lava 'e Siale 'o taa'i Ø e fefinē.
tns pos. erg Siale compl hit abs the woman
"Siale can hit the woman."

Notice that the object, Ø e fefinē, cannot be raised to the higher clause, even though it satisfies both definitions of grammatical subject when Raising applies:

(112) *'oku lava Ø e fefinē 'o taa'i 'e Siale.
tns pos. abs the wom. compl hit erg Siale

The fact that Raising refers to logical subject is substantiated by the informant,[30] whose first reaction to (112) was apparently that its English translation was backwards. While the case marking of (112) can be adjusted to give "The woman hit Siale," the original interpretation was considered to be impossible.

We might try to explain these facts by assuming that Raising has been reordered so that it precedes the passive rule in Tongan. If this were true, Raising would apply when grammatical and logical subject were the same, and there would be no reason to suppose that it was substantially different from (T5a-b). But this solution does not predict the right case marking for (111), where the actor is marked with 'e and thus must have undergone the passive. (113) shows that this is impossible if the actor has been raised before the passive applies:

(113) *'oku lava Ø ('a) Siale 'o taa'i (h)e fefinē.
tns pos. ab̄s prs Siale compl hit the woman

There are two ways out of this dilemma, both of them unfavorable to the H/H hypothesis. We could assume that the ergative is still derived by the passive rule, but that the accusative-to-ergative drift also causes changes in other transformations. In this analysis, the drift has a much more radical effect than that originally predicted by H/H. Or, we could assume that the SD of Raising is not reanalyzed, but that the ergative is produced by a rule other than the Passive:

(T6) Ergative Marking

$$\begin{array}{ccccl} \text{Tns} & \text{V} & \langle \text{NP} \rangle & \text{NP} & \longrightarrow 1 \quad 2 \langle \underline{e} + 3 \rangle\ \underline{\emptyset} + 4 \\ 1 & 2 & 3 & 4 & \end{array}$$

This alternative amounts to a rejection of H/H, but it has the advantage of coping with several of the problems that I have mentioned. If Ergative Marking is responsible for ergative case, then it follows naturally that ergative languages should have a VSO order. Moreover, since Ergative Marking does not reorder actor and object, we can generalize one Raising rule (T5b) for all Polynesian languages. These simplifications provide strong evidence in favor of this alternative.

2.4.2 *Clitic Placement*. The passive also interacts with low-level movement rules. For instance, Pukapukan has a rule of Clitic Placement which moves pronouns to the 'second position' of their clause, i.e., following the tense marker and preceding the verb. Clitic Placement is optional, and derives (115) from (114):

(114) na kalo - wia 0 te wawine e ku.
tns see pass nom the woman agt me

"The woman was seen by me."

(115) na a ku kalo - wia 0 te wawine.
tns prs me see pass nom the woman

"The woman was seen by me."

The rule has a very restricted domain, since it applies only to pronominal actors of passive verbs. Actors of accusative or ergative verbs are not affected, as shown by the following examples:

(116) na vayi 0 te taane i te kovi lewu.
tns strike nom the man acc the pers small

"The man struck the child."

But:

(117) *na a ana vayi i te kovi lewu.
tns pers he struck acc the pers small

(118) na vayi Ø te kovi lewu e te taane.
tns strike abs the prs small erg the man

"The man struck the child."

But:

(119) *na a ana vayi Ø te kovi lewu.
tns pers he strike abs the pers small

Moreover, subjects of intransitives may not undergo the rule:

(120) na kake Ø au ki lunga o te payii.
tns climb nom I to top poss the boat

"I climbed on top of the boat."

But:

(121) *na a ku kake ki lunga o te payii.
tns prs I climb to top poss the boat

(122-4) show that objects are not subject to clitic placement, whatever their particular construction:

(122) *na a ana vayi Ø te taane.
tns pers he strike nom the man

(123) *na a ana kalo - wia e ku.
tns pers she see pass agt me

(124) *na a ana vayi e te taane.
tns pers he strike erg the man

Since the ergative occupies a special place in Pukapukan (§1.3), these restrictions could be described in one of two ways. We could assume that Clitic Placement does not apply to the ergative because its SD includes a passive suffix:

(T7a) Clitic Placement I (optional)[31]

$$\underset{1}{\text{Tns}}\quad \underbrace{\text{V }\underline{\text{(C)(i)a}}}_{2}\quad \underset{3}{\text{X}}\quad \underset{4}{\underline{\text{e}}}\quad \underset{5}{[\text{+pro}]}\quad \underset{6}{\text{Y}} \longrightarrow 1\ \ 5\ \ 2\ \ 3\ \ 6$$

In this case, the ergative is excluded by ordering Clitic Placement after the -(C)(i)a Deletion rule. (Recall that Deletion is optional in Pukapukan, where it produces ergative constructions from passive ones.)

Or we can assume that ergatives are marked in the lexicon for not undergoing Clitic Placement. This view is motivated by the fact that ergatives are exceptions to other rules (§1.3), and allows us to state the SD of Clitic Placement without a passive suffix:

T7b) Clitic Placement II (optional)

$$\underset{1}{\text{Tns}}\quad \underbrace{\text{V}\ \ }_{2}\quad \underset{3}{\text{X}}\quad \underset{4}{\underline{\text{e}}\ [\text{+pro}]} \longrightarrow 1\ \ 4\ \ 2$$

Both of these descriptions must limit the rule to pronouns preceded by e, as otherwise it might apply to subjects of intransitives or actors of accusative verbs. The H/H proposal suggests that this restriction should also appear in ergative languages. This means that CPII will apply in an ergative pattern, since it may affect no NP's besides the actors of transitive verbs. CPI, on the other hand, will not apply at all, because it is ordered after -(C)(i)a Deletion (which is obligatory in ergative languages).

However, some ergative languages have a clitic rule which does not operate in either of these ways. Clitic

Placement in Samoan, for instance, affects the actors of transitive verbs:

(125) na sasa -(ina) e ia Ø le tama.
tns spank pass erg her abs the boy
"The boy was spanked by her."

(126) na ia sasa -(ina) Ø le tama.
tns she spank pass abs the boy
"The boy was spanked by her."

It does not apply to the objects of transitives:

(127) *na ia sasa -(ina) e le fafine.
tns he spank pass erg the woman
("He was spanked by the woman.")

But it does apply to subjects of intransitive verbs, even though they are never preceded by e:

(128) na sau Ø 'oia. "He came."
tns come abs he

(129) na ia sau. "He came."
tns he come

These examples cannot be explained by reordering Clitic Placement before the Passive, because the passive introduces an e on which the other rule depends. Moreover, if Clitic Placement applied first, it would be impossible to generate the passive suffix in (126).

We must therefore conclude that Clitic Placement has been reanalyzed in Samoan, as CPIII

(T7c) Clitic Placement III (optional)

$$\begin{array}{cccc} \text{Tns} & \text{V} & \text{Case} & \text{NP} \\ 1 & 2 & 3 & 4 \end{array} \longrightarrow 1\ 4\ 2$$

or possibly CP IV:

(T7d) Clitic Placement IV (optional)

Tns	V	(Case NP)	Case	NP	→ 1 5 2 3
1	2	3	4	5	

These alternatives correspond to different analyses of the ergative, which may be derived by the Passive rule (for CPIV) or by a separate rule of Ergative Marking. As mentioned above, the first analysis deviates less from the H/H proposal, but the second seems to be closer to the facts. Since Samoan also shows a preference for VSO order, it is simpler to assume that the ergative is independent of the passive, and clitic placement has been reanalyzed as CPIII.

2.5 Double Case Marking. A final example of rule interaction involves the passive and Accusative Marking. This example is particularly interesting because the two rules in question are supposed to be mutually exclusive. In a language like Maori, passive sentences are never subject to accusative marking because they have in some sense already been marked for case.[32] H/H predict that this situation should be preserved in ergative languages; in other words, that all transitive sentences should be passive (= ergative) or accusative, but not both.

But there is at least one Polynesian language which violates this prediction. In Pukapukan, the actor of a transitive sentence may be marked with e at the same time as the object is marked with an accusative i. This double case marking is found only in complement sentences, where it is considered to be a variant of the ergative:

(130) na waka - pono Ø laatou ke tutuli mai i te
tns caus right nom they tns send here acc the

tama e laatou.
boy erg they

"They decided that they
should send the/that boy here."

(131) shows that this construction is ungrammatical elsewhere:

(131) *na tutuli mai i te tama e laatou.
tns send here acc the boy erg them

Double case marking usually indicates that the object is definite, as shown by the translation of (130). It thus contrasts with ordinary ergative case marking, which is allowed in complements but does not have this implied meaning:

(132) na waka - pono Ø laatou ke tutuli mai Ø te
tns caus right nom they tns send here abs the

tama e laatou.
boy erg they

"They decided that they
should send a/the boy here."

In addition, double case marking is restricted on a syntactic level. Like some accusative sentences (§1.3), it does not allow scrambling of actor and object:

(133) *na waka - pono Ø laatou ke tutuli mai e
tns caus right nom they tns send here erg

laatou i te tama.
they acc the boy

It also does not allow the verb to take the passive suffix, although all verbs in the ergative may appear in this construction:

(134) na tala Ø au ki a Yina ke yuyuke -(*ina)
tns say nom I to prs Yina tns open pass

i te puu e na.
acc the dr. erg her

"I told Yina that she should open the/that door."

(134) should be compared with these simple sentence examples:

(135) na yuyuke Ø te puu e na. "She opened the door."
tns open abs the door erg her

(136) na yuyuke -ina Ø te puu e na.
tns open pass nom the door agt her

"The door was opened by her."

Finally, it is important that the i in (134) must be the accusative marker, and not some other marker which indicates specificity or definiteness. This fact is established by the informant, who stated that (134) is incorrect because the accusative i and the passive suffix never co-occur.

Double case marking provides another counterexample to H/H, who apparently assume that the two rules of Passive and Accusative Marking will never interact. The counterexample is even stronger because of the restrictions on the double case marking rule. For suppose that we wanted to account for (130) within the H/H proposal (and the proposal was modified so that rule interaction was allowed). We would probably rewrite Accusative Marking or the Passive so that one could apply to the output of the other. This interaction could take place only in complement sentences, and would have to be restricted so that sentences like (133) would not be derived.

But notice what it is that these restrictions do. They state that passive sentences may undergo scrambling, but sentences with double case marking may not; that passive verbs can take the passive suffix, but verbs in the double case marking construction cannot --in short, that the only feature which the passive shares with double case marking is that both mark the actor NP with _e_. Under the circumstances, we can reasonably conclude that (130) is not derived by the Passive Rule. It is produced instead by a Double Case Marking rule, which marks the actor of a complement with _e_ and the object with _i_:

(T8) Double Case Marking (optional)[33]

$$\underbrace{X\quad Tns\quad V}_{1}\quad \underset{2}{NP}\quad \underset{3}{NP}\quad \underset{4}{Y} \longrightarrow 1\quad \underline{i}+3\quad \underline{e}+2\quad 4$$

$$\text{if } X \neq \#$$

Since this rule is distinct from the passive but generates sequences of _e_ NP, it contradicts the basic premise of H/H: that all ergative case marking in Polynesian is derived by the passive rule.

2.6 An Alternative Proposal. It thus appears that the H/H proposal is too strong to account for the facts of Polynesian. I have tried to show that ergative languages have several characteristics in common; they are rigidly VSO, they have reanalyzed the passive suffix, and they have certain rules which refer to logical subject (but do not in accusative languages). As these characteristics are not found in accusative languages, we cannot say that an ergative language is just an accusative one in which the passive is required. Apparently, the difference between the accusative and ergative types extends beyond any one rule, even when this rule is as functionally important as the passive.

Still, the evidence does point to some kind of accusative-to-ergative drift. There is clearly a historical tendency for the passive to be favored in Polynesian languages. Moreover, all ergative languages have morphemes which can be traced to the passive, although they differ from it in failing to function as a single rule. For example, Tongan has a passive suffix, but this can occur in sentences with accusative case marking; Samoan has 'passive-like' case markers, but these are independent of -(C)(i)a.

These facts suggest an alternative proposal to H/H when we remember that the passive has a radical effect on sentence structure. In order to shift the underlying relations of subject and object, the passive makes use of three different structural changes: these are agent extraposition, subject formation (= case marking), and attaching the passive suffix.

Historically, it is not unreasonable that the frequency of the passive should come to conflict with its status as a radical rule. If the passive continued to apply more often over time, it might become unclear to new speakers of a language whether it was a derived construction at all. Children might then assume that it was not the product of a radical rule, but that it correctly reflected the relations of underlying structure. This assumption obviously amounts to a reanalysis of the passive. Under the new analysis, passive sentences would be treated as the products of three low level (and therefore unimportant) transformations: ergative case marking, scrambling, and -(C)(i)a suffixation. These would often co-occur, but they would be essentially unrelated, and thus would also be allowed to apply by themselves. It is significant that the results of this reanalysis would be exactly like the rules which we find in Tongan and Samoan.

This proposal has several clear advantages over the one suggested by H/H. First, it explains why ergative languages should have rules which resemble the passive, but are not the same on a formal level. Second, it accounts for later changes in these rules which are difficult to explain as changes in the passive. We can account for the fact that ergative languages are VSO by saying that they have eliminated the scrambling rule (except in special cases); we can explain the extension of -(C)(i)a by saying that the suffix attachment rule has been generalized. These explanations make sense in a system where the rules are separate and can undergo historical changes as such; they are less satisfactory when they are applied to dependent parts of the passive rule.

In addition, the proposal has two rather interesting consequences. One is that there is a threshold beyond which a radical syntactic rule cannot apply, without obscuring the relations of underlying structure. If an optional rule like the passive[34] should happen to cross this threshold, it would be subject to the sort of reanalysis described above. I believe that this historical fact is motivated by a more general constraint on language, that unmarked constructions should reflect their underlying structures in a more or less transparent way. This principle is analogous to one suggested independently by Kiparsky (1971), in a discussion of phonological and paradigmatic change.

Another consequence is that the reanalysis of the passive must occur towards the beginning of the accusative-to-ergative drift. For instance, Pukapukan seems to be a basically accusative language, since its ergative construction is new and restricted to casual speech. Most examples of this construction can be derived by the H/H proposal in an awkward, but reasonably adequate way (§1.3). But Pukapukan also has a double case marking construction, which cannot be derived by the passive rule. Since the appearance of double case marking suggests that the passive has already been reanalyzed, we might guess that no ergative constructions in Pukapukan are derived as H/H describe. It is tempting to conclude from this that the reanalysis occurs before the ergative is introduced into a language.

Both the proposal and its consequences deal with language change, but they are based on one synchronic and controversial assumption. This is that case marking in ergative languages is not arbitrary and superficial, but reflects the grammatical relations of some level of underlying structure.

This assumption is directly opposed to that of H/H, who claim that an accusative deep structure is basic, and ergative case marking is derived. However, there are reasons which suggest that it is preferable to the H/H view.

As we have seen, H/H claim that ergative case marking is derived by means of the passive rule. Since the passive is a radical rule, but Accusative Marking is not, this means that ergative constructions should always be marked with respect to accusative ones. However, this is not always true in Polynesian. In Tongan and Samoan, the ergative is normal for the majority of transitive verbs, but the accusative is limited to a small and clearly defined class. There are also instances of accusative verbs which come to be used in the ergative construction, but no examples of the opposite change. Samoan *va'ai* 'to look at' is identified as accusative by Churchward (1928), but it is accepted in the ergative construction by Johnson and Harmon (1962), and my informants.[35]

(137)	'ua	'uma	'ona	va'ai	*Ø*	laa'ua	*'i*	le	fafine.
	tns	finish	compl	look	nom	they	acc	the	woman

"They (du.) have finished examining the woman."

(138)	'ua	'uma	'ona	va'ai	*e*	laa'ua	*Ø*	le	fafine.
	tns	finish	compl	look	erg	they	abs	the	woman

"They (du.) have finished looking at the woman."

On the other hand, an ergative verb like *fasi* 'to hit', can never take accusative marking:

(139) 'ua fasi e le teine 0 le tama.
tns hit erg the girl abs the boy

"The girl hit the boy."

(140) *'ua fasi 0 le teine 'i le tama.
tns hit nom the girl acc the boy

Examples like this establish that, on an absolute level, ergative case marking is not less arbitrary than the accusative type. Nonetheless, it is also true that ergative relations cannot be a part of deep structure. I have mentioned that both Tongan and Samoan have rules which refer to logical subject; that is, the actor of a transitive verb and the subject of an intransitive. If we were to assign an ergative deep structure to languages like these, we would not be able to capture the fact that some rules apply in this other pattern.

I believe that this complicated problem can be solved by assuming that not one, but two types of relations are involved. These are the deep structure relations of logical subject and object, and the more superficial relations which are expressed by case. My claim makes a distinction between logical subject and the nominative case, although this distinction is usually not explicit in previous work.[36] Nonetheless, it is attractive for several reasons.

First, it accounts for the fact that both accusative and ergative languages have rules which refer to logical subject. (It is this relation which is obviously the candidate for universal grammar.) Second, it suggests that accusative and ergative are notions which are relevant at some point in intermediate structure. This suggestion explains why ergative case marking should be unmarked in some languages but not others; and why an accusative-to-ergative drift should be a possible language change.[37]

As I have mentioned, a counterproposal to this would claim that accusative case marking is basic, and ergative case marking is arbitrary and derived. It is interesting to conside what evidence would lead us to reject this counterproposal. One type of evidence might establish that many transformations in Polynesian behave essentially like case marking rules, in that they differ systematically from accusative to ergative languages. This demonstration would make clear that accusative and ergative case marking are equally derived, and would also suggest that the rules which derive them are not arbitrary (although they are low-level). Instead, they would be one of a whole complex of rules which would conform to the organizing principles of accusative and ergative types of case. These principles would be directly reflected by (but not based on) the case marking of simple sentences.

In the following sections, I will attempt such a demonstration for Polynesian. I will examine two types of rules which classify NP's --clause-restricted rules and two-clause rules-- and discuss their structures in accusative and ergative languages. The discussion suggests that these rules differ systematically between languages, and that the differences are based on a well-defined relationship between case marking and the other rules.

3.0 Case Marking and Clause Restricted Rules

The next two sections describe rules whose domains are restricted to simple sentences (or clauses). Clitic Placement moves a pronoun to a position following the tense marker of its clause (second position in the sentence). Nominalization is a set of rules which turn complement sentences into derived NP's, if the clauses are immediately dominated by the node NP in underlying structure. Since these processes resemble case marking rules in being clause-restricted, we might expect their case structure to be affected by the accusative-to-ergative drift. This expectation turns out to be true, as is demonstrated in §4-5.

4.0 Nominalization

The four languages described in §1 all have a process of nominalization, which can be reconstructed for proto-Polynesian. Although the details of this process vary from language to language, there are four rules which are apparently common to all: these are Nominalization Formation, Extraposition, Possessive Marking and Possessive Preposing. Nominwlization Formation makes the verb of a complement into a noun, by giving it a nominalizing suffix and replacing its tense marker with an article. Extraposition may optionally reorder the actor and object of the nominalized verb. Possessive Marking places the NP next to the verb into the possessive, after Extraposition has applied. Finally, the optional rule of Possessive Preposing moves certain nouns to the second position in the nominalization.

The four rules are found in all of the representative languages, but they have different conditions, and are ordered differently, in each. (The preceding paragraph describes the rules as they can be reconstructed for proto-Polynesian.) As I try to show, these differences are not arbitrary, but are usually designed to give the nominalizations a specific case relation. The nominalization rules are therefore affected by the accusative-to-ergative drift, and the discussion gives some idea of the complicated ways in which this influence is brought about.

4.0.1 Preliminaries. One fact which will be especially relevant in §4.1-5 is the way that possessives are formed in Polynesian. Most Polynesian languages[38] make a distinction between alienable and inalienable possession. Alienable possession is marked by a particle a which precedes the possessor NP; inalienable possession is marked by a particle o. Consider these examples from Maori:

(141) te pukapuka a te wahine "the book of/owned
the book alien. the woman by the woman"

(142) te ihu o te tangata "the nose of the
the nose inal. the man man"

If the possessor is a full noun it follows the possessed noun. It follows the article if it is a clitic:

(143) t - *a* - na pukapuka "her book/the book
the alien. her book owned by her"

(144) t - *o* - na ihu "his nose"
the inal. him nose

Most possessed nouns are typically found with one of the *a* or *o* possessive, but the other construction is usually possible given the proper semantic circumstances. The resulting contrast is often related to that between the subject and object of a sentence (Williams,1965; S. Churchward, 1951; C. Churchward, 1953). In (141), for example, *wahine* is the active, subject-like possessor of *pukapuka*, but in (145), it is passive and object-like:

(145) te pukapuka *o* te wahine "the book of/about
the book inalien. the woman the woman"

4.1 *Maori*. Nominalizations in Maori are practically a direct reflex of the type that we can reconstruct for proto-Polynesian. Basically, they are derived from clauses by a rule of Nominalization Formation, which replaces the tense marker with an article and adds a nominalizing suffix to the verb:

(TM1) Nominalization Formation

$$\begin{array}{cccccccccccc} [_{NP} & [_{S} & Tns & V & X & _{S}] & _{NP}] & \longrightarrow & 1 & Art & 4 + \underline{(C)(a)nga} \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & & & 5 & 7 \end{array}$$

The suffix has the form -*(C)(a)nga* in all Polynesian languages, and its initial -C is the same consonant that appears in the passive suffix (§1.1). As we might expect, -(C)(i)a and -(C)(a)nga are subject to the same types of historical change, including development of an unmarked form and semantic differentiation of the -C's (§2.2). For instance, Maori uses the verb-specific variants of -(C)(a)nga, but it also has an unmarked form -tanga:

(145) te haere - nga "(someone's) going"
the go nomin

(146) te haere - tanga "(someone's) going"
the go nomin

(145-6) are derived from (147):

(147) ka haere. "(Someone) goes."
tns go

The NP closest to the nominalized verb is usually placed in the possessive by a rule of Possessive Marking.[39] Subjects of intransitives appear in the o construction:

(148) ko te haere - nga o Hoone, he mea kino.
*a
top the go nomin poss John a thing bad
"John's going, [it's] a bad thing."

Compare:

(149) ka haere 0 Hoone. "John goes."
tns go nom John

Actors of transitives appear in the a construction:

(150) kua whakawaa - tia 0 te patu - nga a Hoone
*o
tns investig. pass nom the hit nomin poss John
i te wahine.
acc the woman

"John's killing the woman
has been investigated."

From:

(151) i patu _0_ Hoone i te wahine.
tns hit nom John acc the woman

"John killed the woman."

In these cases Possessive Marking is apparently obligatory, and must affect the NP closest to the verb. (152-3) show that a nominalization is ungrammatical when Possessive Marking has not applied:

(152) *ko te haere - nga _0_ Hoone, he mea kino.
top the go nomin nom John a thing bad

(153) *kua whakawaa - tia 0 te patu - nga _0_ Hoone
tns investig. pass nom the hit nomin nom John

i te wahine.
acc the woman

(154) shows what happens when the object of a transitive is nominalized instead:

(154) *kua whakawaa - tia 0 te patu - nga 0 Hoone
tns investig. pass nom the hit nomin nom John

_o_ te wahine.
_a_
poss the woman

It is important that the choice of _a_ or _o_ in these examples is determined exclusively by grammatical relations. Nominalizations are therefore different from ordinary possessed nouns, whose choice of _a_ or _o_ is partly affected by semantics.

Possessive Marking applies after a rule of Extraposition, which may optionally reorder the actor and object of a nominalization. Extraposition does not affect the object at all, but extraposes the actor and marks it with the particle e:

(155) kua whakawaa - tia Ø te patu - nga i te
tns investig. pass nom the hit nomin acc the

wahine e Hoone.
*Ø
woman agt John

"The killing [of] the woman by John has been investigated."

When Possessive Marking applies to (155), it is the object which is placed in the possessive, and takes the o construction. (156) shows that the a construction is not allowed instead:

(156) kua whakawaa - tia Ø te patu - nga o te
*a
tns investig. pass nom the hit nomin poss the

wahine e Hoone.
woman agt John

"The killing of the woman by John has been investigated."

Notice that Possessive Marking is not required when the object is closest to the verb, since (155) was declared by the informant to be acceptable, although not as good as (156). On the other hand, the rule is obligatory when it applies to actors or subjects. The reason for this difference is not entirely clear to me, but seems to come from a general requirement that all nouns be preceded by overt particles in nominalizations. (This amounts to saying that a complex sentence should have only one morphological subject, since the nominative is the only case which is not marked by an overt particle in Maori.[40]) As actors and subjects are preceded by Ø when Possessive Marking applies, they are obliged to undergo the rule. Objects, however, are already marked by the particle i, so Possessive Marking is optional for them.

These facts suggest that Possessive Marking should be formalized as follows:

(TM3) Possessive Marking

$$[_{NP}\ \underbrace{X\ \ \underline{(C)(a)nga}}_{1}\ \ \underset{2}{\text{Case}}\ \ \underbrace{NP \langle \underline{i}\ NP \rangle\ Y}_{3}\ {}_{NP}]$$

$$\longrightarrow 1 \begin{Bmatrix} \langle \underline{a} \rangle \\ \underline{o} \end{Bmatrix} 3 \qquad \text{oblig. if } 2 = \underline{\text{Ø}} \quad \text{optional otherwise}$$

Here the angled brackets indicate that the two parts of the rule are disjunctive, as double possessives are apparently not allowed.[41]

The Extraposition rule is slightly more difficult to formalize. Since Extraposition reorders the NP's and introduces the particle e, we might be tempted to consider it as a variant of the passive rule. Sentences like (156) might then be derived from passive sentences which would undergo Nominalization Formation and Possessive Marking, as well as a -(C)(i)a Deletion rule. However, this analysis cannot be right, because Maori has a separate set of passive nominalizations:

(157) he kino Ø te epa - ina - tanga o te
*a
a bad nom the throw pass nomin poss the

kupenga e te tangata.
net agt the man

"The net's being thrown by
the man is a bad [thing]."

The nominalization in (157) is obviously derived from (158):

(158) ka epa - ina Ø te kupenga e te tangata.
tns throw pass nom the net agt the man

"The net will be thrown
by the man."

Like (156), a passive nominalization places its object in the o possessive. But it differs from (156) in not allowing the actor to be possessed. Compare (150) with the following:

(159) *he kino Ø te epa - ina- tanga a te tangata
a bad nom the throw pass nomin poss the man

o te kupenga.
Ø
i
the net

Moreover, it may not occur in the double case marking form of (155):

(160) *he kino Ø te epa - ina - tanga i te kupenga
Ø
a bad nom the throw pass nomin the net

e te tangata.
agt the man

And it must always take Possessive Marking, as a result (v. 160).

These characteristics are so opposed to those of (156) that both types of nominalizations could not be derived from the same source. Rather, (156) must come from an accusative sentence, with Extraposition and Possessive Marking operating as I have described. Extraposition thus resembles a subpart of the passive, which operates on nominalizations much like Agent Postposing does in English (Chomsky, 1970). We can formalize the Maori rule as:

(TM2) Agent Extraposition (optional)

$$[_{NP}\ \underbrace{X\ \ \underline{(C)(a)nga}}_{1}\ \ \underbrace{\underline{Ø}}_{2}\ \ \underbrace{NP}_{3}\ \ \underbrace{\underline{i}\ \ NP}_{4}\ \ \underbrace{Y\ {}_{NP}]}_{5}$$

$$\longrightarrow\ 1\ \ 4\ \ \underline{e}\ \ 3\ \ 5$$

Extraposition is ordered after Nominalization Formation, but before Possessive Marking.

Finally, the possessor may be preposed to 'second position' in the nominalization, after the article and before the nominalized verb. This rule of Possessive Preposing can be reconstructed for proto-Polynesian, where it was optional for pronouns in all possessive constructions (not just nominalizations).[42] In Maori, however, the rule has been extended so that it is allowed for all possessive nouns, but required for pronouns. Compare (148, 150, 156, 157) with the following examples:

(161) ko t - $\underline{o}$ Hoone haere - nga, he mea kino.
*$\overline{a}$
top the poss J. go nomin a thing bad
"John's going, [it's] a bad thing."

(162) kua whakawaa - tia Ø t - $\underline{a}$ - na patu - nga
*$\overline{o}$
tns investig. pass nom the poss him kill nomin
i te wahine.
acc the woman

"His killing the woman
has been investigated."

(163) kua whakawaa - tia Ø t - o te wahine
*a
tns investig. pass nom the poss the woman

patu - (a) - tanga e Hoone.
kill pass nomin agt John

"The woman's killing (being
killed) by John has been
investigated."

These examples of ordinary possessive constructions should be compared with (141, 145):

(164) t - a te wahine pukapuka "the book of/owned
the poss the woman book by the woman"

(165) t - o te wahine pukapuka "the book of/about
the poss the woman book the woman"

Notice that Preposing does not affect the choice of an a or o construction, and so can be formalized as:

(TM4) Possessive Preposing (optional)

$$\underbrace{[_{NP}\ \text{Art}}_{1}\quad \underbrace{\text{V}\ \underline{\text{(C)(a)nga}}}_{2}\quad \underbrace{\text{Poss NP}}_{3}\quad \underbrace{\text{X}\ _{NP}]}_{4}$$

$$\longrightarrow 1\ \ 3\ \ 2\ \ 4$$

and ordered after Possessive Marking.

To summarize, I list the nominalization rules for Maori in the order in which they apply.

Possessive Formation

Agent Extraposition (optional)

Possessive Marking

Possessive Preposing (optional)

It is particularly important for our purposes that Possessive Marking has an ergative structure in Maori. The rule treats the subject of an intransitive like the object of a transitive verb, since it places both NP's in the *o* form of the possessive. This form has been identified previously with the notion of semantic object (§4.0.1). On the other hand, the actor of a transitive appears in the *a* possessive, and is the only member of a nominalization for which this construction is allowed.

4.2 *Tongan*. The details of nominalization in Tongan are rather different. Nominalizations are still formed from complement sentences by replacing the tense marker with an article, but in Tongan a nominalizing suffix is not attached to the verb. The nominalization in (166), for example, is productively derived from (167):

(166) 'i he 'osi 0 'e - nau akō, pea nau kai.
at the end abs poss they study then they eat

"At the ending [of] their studies, they ate."

(167) na'e 'osi 0 'e - nau akō. "Their studies ended."
tns end abs poss they study

And the variants of the suffix -(C)(a)nga are all ungrammatical:

(168) *'i he 'osi - 'anga 0 'e - nau akō, pea nau kai.
at the end nomin abs poss they study then they eat

This fact simplifies the rule of Nominalization Formation, which we can write as follows:

(TT1) Nominalization Formation

$$\begin{array}{cccccc} [_{NP} & [_{S} & Tns & X & _{S}] & _{NP}] \\ 1 & 2 & 3 & 4 & 5 & 6 \end{array} \longrightarrow 1 \quad Art \quad 4 \quad 6$$

Compare this rule with Nominalization Formation in Maori.

Although it is no longer used in nominalizations,[43] it is interesting that -*(C)(a)nga* does survive in a secondary function in Tongan. Churchward (1953) states that -*(C)(a)nga* is a noun-forming suffix which indicates "the place, the seat, the source, or the means, of the action or state specified."[44] This type of lexical derivation is limited to a small class of verbs (mostly intransitives):

(169)	ako	'learn'	ako-'anga	'educational institution'
	tu'u	'stand'	tu'u-'anga	'standing place'
But:	'alu	'go'	*'alu-'anga	

It also has slightly different semantic consequences, depending on which variant of the suffix is used. For instance, -*'anga* is used to form place nouns, but -*nga* "indicates thing rather than place:"[45]

(170)	ako	'learn'	aako-nga	'student'
	tu'u	'stand'	tu'u-nga	'ladder; standing, status'

-*(C)(a)nga* is thus very similar to the passive suffix in Tongan, which does not form passive sentences, but is used in lexical derivation. On the other hand, -*(C)(i)a* and -*(C)(a)nga* are both productive in Maori. These resemblances provide some slight evidence that the two suffixes have similar histories in Polynesian. If so, it would not be unreasonable to claim that -(C)(a)nga is also affected by the accusative-to-ergative drift.

The nominalizations then undergo a rule of Possessive Marking, which is required if the NP closest to the nominalized verb is a pronoun. Possessive Marking places the subject (of an intransitive) or the actor (of a transitive) in the 'a construction. Since scrambling has not yet applied, an object will never be subject to the rule:[46]

(TT2) Subject Possessive Marking (first try)

$$[_{NP}\underbrace{\text{Art}\quad\text{V}}_{1}\quad\underset{2}{\text{Case}}\quad\underbrace{\text{NP}\quad\text{X}}_{3}\,_{NP}] \longrightarrow 1\ \text{'}\underline{\text{a}}\ \ 3$$

(171) shows Possessive Marking applying to a pronominal subject. (The possessive pronoun is then moved forward by the Possessive Preposing rule.)

(171) 'oku fai Ø 'e - ku huo.[47] "My hoeing is being done."
*ho
tns do abs poss my hoe

From:

(172) 'oku ou huo. "I am hoeing."
tns I hoe

(173) shows the rule applying to a pronominal actor:

(173) 'oku 'ikai tonu Ø 'e - ne taki Ø e fonuā.
*ho
tns not good abs poss him lead abs the land
"His leading the land is not good."

From:

(174) 'oku ne taki 0 e fonuā. "He was leading the land."
tns he lead abs the land

On the other hand, Possessive Marking is optional when the subject of an intransitive is not a pronoun.[48] The NP in (175) has presumably undergone the rule, since its 'a is not omitted, even in casual speech:

(175) 'oku tuai 0 e 'alu 'a e tamasi'ī.
*'o
tns slow abs the go poss the child

"The going of the child is slow."

From:

(176) na'e 'alu 0 e tamasi'ī. "The child went."
tns go abs the child

But the NP in (177) has the case marking of a simple sentence:

(177) 'oku faka-mamahi 'aupito 0 e si'i mate
tns caus. sad very abs the-emot die

0 e fa'eē mo e si'i faanaū.
abs the moth. and the-emot child.

"The unfortunate dying [of] the mother and the poor children is very sad."

From:

(178) na'e mate 0 e fa'eē mo e si'i faanaū.
tns die abs the moth. and the-emot child.

"The mother and the poor children died."

The Tongan version of Possessive Marking therefore differs from Maori, where the rule was required for any NP preceded by the particle 0. We might stop for a moment to consider why this difference should exist. I have suggested that Possessive Marking has the conditions which it does in

Maori because the language formally recognizes only one subject in a complex sentence. This subject is placed in the nominative (marked by Ø), which directly reflects the relation of logical subject. Now Tongan also has a 'nominative' (marked by Ø), but this does not refer to logical subject because of the accusative-to-ergative drift. Since the language has no way to formalize the subject relation on a morphological level, it seems to have dropped the requirement that subjects undergo Possessive Marking.

Possessive Marking is never allowed when the actor of a transitive is a full noun. Contrast (179), which is ungrammatical, with the Maori example (150):

(179) *'oku 'ikai tonu Ø e taki 'a/'o e tu'ī
tns not good abs the lead poss the king
Ø e fonuā.
abs the land

(179) is obviously derived from the simple sentence (180). (181) gives the correct form of the nominalization:

(180) 'oku taki 'e he tu'ī Ø e fonuā.
tns lead erg the king abs the land
"The king was leading the land."

(181) 'oku 'ikai tonu Ø e taki 'e he tu'ī
tns not good abs the lead erg the king
Ø e fonuā.
abs the land
"The king's leading the land is not good."

The fact that Possessive Marking is not obligatory here can be explained by the overt e which precedes the actor NP.[49] (Recall that in Maori, Possessive Marking was optional for objects, which were preceded by an overt case particle i.) It is more surprising that the rule cannot apply at all. We can guess that this restriction might have been designed to avoid a possible confusion of the possessive and absolutive 'a's. Since 'a was until recently the absolutive marker in Tongan, sentences like (182) might have been hopelessly ambiguous, and thus excluded from the grammar:

(182) *'oku 'ikai tonu 'a e taki 'a e tu'ī 'a
tns not good abs the lead poss the king abs

e fonuā.
the land

On the other hand, the restriction might have been developed because the possessive 'a and the ergative e have basically the same function in Tongan. The possessive 'a places a possessor in an active, agent-like relation to the noun; the ergative 'e creates the same relation between an actor and its transitive verb. As the possessor and actor both occupy the same position in the sentence, a rule which just replaces 'e with 'a would be functionally useless. We can imagine that this fact might keep Possessive Marking from applying to full noun actors at all.

Both of these explanations agree with the fact that Possessive Marking must apply to actors and subjects which are pronouns. This is because all pronoun actors and subjects must undergo a preposing rule, which separates them from the other NP's of the sentence (v. 171-4). Sentences (183-4) demonstrate that preposing is limited to pronouns:

(183) *'oku 'ikai tonu Ø e - 'a e tu'ī taki
tns not good abs the poss the king lead
Ø e fonuā.
abs the land

(184) *'oku tuai Ø e - 'a e tamasi'ī 'alu.
tns slow abs the poss the child go

Nonetheless, there is some evidence from other languages which suggests that the second explanation is closer to the truth. Samoan, for example, has a rule of Possessive Marking which applies to pronominal actors, but only marginally to actors which are full nouns. As Samoan has never had an absolutive marker 'a, this restriction cannot be explained by appealing to homophony.[50]

The complete Possessive Marking rule is therefore as follows:

(TT2) Subject Possessive Marking

$$[_{NP}\underbrace{\text{Art}\quad \text{V}}_{1}\ \ \underset{2}{\text{Case}}\ \ \underset{3}{\text{NP}}\ \ \underbrace{\text{X}\ {}_{NP}]}_{4} \longrightarrow 1\ \text{'}\underline{\text{a}}\ \ 3\ \ 4$$

oblig. if 3 = [+pro]
otherwise, opt. if
2 = $\underline{Ø}$

Nominalizations which are transitive and have not undergone this rule are subject to a rule of Extraposition. This optional rule reorders the actor and object of a verb, but seems to have no other effect:

(TT3) Agent Extraposition (optional)

$$[_{NP}\underbrace{\text{Art}\quad \text{V}}_{1}\ \ \underbrace{\text{'}\underline{\text{e}}\ \text{NP}}_{2}\ \ \underbrace{\underline{Ø}\ \text{NP}}_{3}\ \ \underbrace{\text{X}\ {}_{NP}]}_{4} \longrightarrow 1\ \ 3\ \ 2\ \ 4$$

Extraposition has applied in (185), but is not allowed in (186), where the actor is in the possessive:[51]

(185) 'oku 'ikai tonu Ø e taki *'o* e fonuā
*'a
tns not good abs the lead poss the land

'e he tu'ī.
erg the king

"The leading of the land
by the king is not good."

Compare:

(186) *'oku 'ikai tonu Ø e taki *'o* e fonuā
'a
tns not good abs the lead poss the land

'e - ne.
poss him

It is interesting that the *'e* in (185) need not be introduced by Extraposition, since *'e* normally appears before the ergative noun in Tongan. The Tongan rule is thus simpler than Extraposition in Maori, although the simplification is admittedly minor.

Objects which have been left closest to the verb by Extraposition must then undergo an Object Possessive Marking rule. Object Possessive Marking is obligatory, and places the object in the *'o* form of the possessive (v. 185).[52] Note that the object must be closest to the verb to undergo the rule:

(187) *'oku 'ikai tonu Ø e taki 'e he tu'ī
tns not good abs the lead erg the king

'o e fonuā.
poss the land

Further, the rule must apply whenever this condition is satisfied:

(188) *'oku 'ikai tonu Ø e taki Ø e fonuā
tns not good abs the lead abs the land

'e he tu'ī.
erg the king

Since Subject and Object Possessive Marking are collapsed as one rule in Maori, it may seem peculiar that they are separated in Tongan. However, this ordering is needed to explain the behavior of nominalizations which can undergo both rules. If a nominalization has an actor which can take Subject Possessive Marking, Extraposition and Object Possessive Marking can never apply:

(189) *'oku 'ikai tonu Ø e taki 'o e fonuā
tns not good abs the lead poss the land

'e ia.
erg him

("The leading of the land
by him is not good.")

This restriction holds even when the object is a pronoun, as illustrated by (190):

(190) *'oku 'ikai tonu Ø ho - no taki 'e ia.
tns not good abs poss its lead erg him

("*Its leading by him is
not good.")

The fact that the 'a possessive dominates over the 'o possessive in this way is captured by ordering Subject Possessive Marking before the Extraposition rule. Then Object Possessive Marking will work as follows:

(TT4) Object Possessive Marking (obligatory)

$$[_{NP}\underbrace{\text{Art}\quad\text{V}}_{1}\quad\underset{2}{\underline{\emptyset}}\quad\underbrace{\text{NP}\ \underline{\text{'e}}\ \text{NP}\quad\text{X}}_{3}\ _{NP}] \longrightarrow 1\quad\underline{\text{'o}}\quad 3$$

Finally, possessor NP's which are pronouns are preposed to second position in the NP. Preposing is obligatory, as in Maori, and applies to all possessive constructions, not just nominalizations:

(191) 'oku 'ikai tonu Ø ho - no taki 'e he tu'ī.
*'e
tns not good abs poss its lead erg the king

"*Its leading by the king is not good."

The preposed pronoun in (191) is the object of the verb. (171) and (173) give examples of the other preposed pronouns.

Preposing is connected to one other fact about nominalizations which seems to be restricted to Tongan. In Tongan, all possessive pronouns may leave emphatic copies behind. These copies are described as 'reflexive' by Churchward (1953), and are illustrated below:

(192) 'oku fai Ø 'e - ku huo 'a - ku.
tns do abs poss my hoe poss my

"My hoeing is being done."

(193) 'oku 'ikai tonu Ø 'e - ne taki 'e ia
tns not good abs poss his lead erg him

Ø e fonuā.
abs the land

"His leading the land is not good."

(194) 'oku 'ikai tonu Ø ho - no taki o'o - na.
tns not good abs poss his lead poss his

"His being led is not good."

Notice that a copied subject or object appears in the possessive, but a copied actor is preceded by the ergative marker 'e (193). This fact can be explained if we assume that the copies are pronouns which act like full nouns (not like clitics). Possessive Preposing then becomes:

(TT5) Possessive Preposing (obligatory)

$$\underbrace{[_{NP}\ \text{Art}}_{1}\quad \underset{2}{\text{V}}\quad \underset{3}{\text{Poss}}\quad \underset{4}{[+\text{pro}]}\quad \underbrace{\langle \underline{0}\ \text{NP}\rangle\ \text{X}\ _{NP}]}_{5}$$

$$\longrightarrow\quad 1\quad 3\quad \overset{4}{[+\text{clit}]}\quad 2\ \begin{Bmatrix} 3 \\ \langle \underline{e} \rangle \end{Bmatrix}\ \overset{4}{[-\text{clit}]}\quad 5$$

The emphatic pronouns may then be deleted by a rule of Emphatic Copy Deletion:

(TT6) Emphatic Copy Deletion (optional)

$$\underset{1}{\text{X}}\ \underbrace{\begin{Bmatrix} \text{Poss} \\ \underline{e} \end{Bmatrix} \begin{bmatrix} +\text{pro} \\ -\text{clit} \end{bmatrix}}_{2}\ \underset{3}{\text{Y}} \longrightarrow 1\quad 3$$

Notice that Possessive Preposing must follow the Object Possessive Marking rule.

It is significant that Emphatic Copy Deletion and Possessive Preposing have parallels elsewhere in Tongan, although they do not operate this way in any of the other representative languages.[53] We will see that emphatic copies are allowed for pronouns in simple sentences which are actors of transitive, or subjects of intransitive verbs.

To summarize, the rules for nominalization in Tongan are:

Nominalization Formation
Subject Possessive Marking
Agent Extraposition (optional)
Object Possessive Marking
Possessive Preposing
Emphatic Copy Deletion (optional)

These rules differ from nominalization in Maori in three main ways: (1) Nominalization Formation does not introduce a nominalizing suffix, (2) a rule of Emphatic Copy Deletion has been added, and (3) Possessive Marking is separated into two rules. The split of Possessive Marking is probably related to the fact that the two rules have an accusative structure in Tongan. As we have seen, Subject Possessive Marking places actors and subjects in the 'a possessive, which is identified with the notion of semantic agent. On the other hand, objects are placed in the 'o possessive. and are the only grammatical NP's for which this construction is allowed.

4.3 Pukapukan. The preceding sections have suggested that the case marking of simple sentences is somehow related to the Possessive Marking rule. Possessive Marking is ergative in an accusative language (Maori), but accusative in an ergative one (Tongan). If this relation is a general one, we would expect Possessive Marking to have a mixed type of case in languages in the middle of the continuum. This prediction turns out to be true for Pukapukan, although its nominalizations are otherwise much like Maori.

Nominalizations in Pukapukan are formed in the usual way, by replacing the tense marker of a complement with an article and adding a suffix -*(C)(a)nga* to the verb. Several variants of the suffix are used, but the most common seems to be -*nga* (or sometimes, -*anga*):

(195) te patu - *nga* i naa toa. "the killing [of] the warriors"
the hit nomin acc the-pl warrior

(195) is derived from (196):

(196) na patu i naa toa. "(Someone) killed the warriors"
tns hit acc the-pl warrior

As the nominalizing suffix is productive, the rule which inserts it is essentially the same as Nominalization Formation in Maori. It can be formalized quite simply:

(TP1) Nominalization Formation

$$\begin{array}{cccccccccccccc} [_{NP} & [_{S} & \text{Tns} & V & X & {}_{S}] & {}_{NP}] & \longrightarrow & 1 & 3 & 4 & + & \underline{(C)(a)nga} \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & & & & & 5 & 7 \end{array}$$

and should be compared with the Maori rule.

The NP closest to the nominalized verb is usually placed in the possessive by a rule of Possessive Marking. This rule is obligatory^{54} for actors of accusative verbs, which take *a*:

(197) e mea kino 0 te patu - nga *a* Turi
**ō*
a thing bad nom the hit nomin poss Turi
i te wawine.
acc the woman

From:

(198) na patu Ø i a Turi i te wawine.
tns hit nom pers Turi acc the woman

"Turi killed the woman."

And it is obligatory for objects of passive verbs, which take o:

(199) e mea kino Ø te patu - a - nga o te wawine
*a
a thing bad nom the hit pass nomin poss the wom.

e Turi.
agt Turi

"The woman's being killed by Turi is a bad thing."

From:

(200) na patu - a Ø te wawine e Turi.
tns hit pass nom the woman agt Turi

"The woman was killed by Turi."

Since both types of NP's are preceded by Ø before the rule applies, these sentences suggest that Pukapukan requires overt particles before NP's in nominalizations. I have suggested that this constraint only operates when the Ø particle marks the nominative, as in Maori. Pukapukan provides more evidence for this claim, because its ergative nominalizations do not undergo the rule. This is, of course, exactly what we would predict if the constraint on overt particles was connected to case marking in some way.

Possessive Marking does not apply to ergative nominalizations at all, as illustrated by (201):

(201) *e mea kino Ø te patu - nga a Turi
ō
a thing bad nom the hit nomin poss Turi

Ø te wawine.
abs the woman

Compare:[55]

(202) na patu e Turi Ø te wawine. "Turi killed
tns hit erg Turi abs the woman the woman."

(203) shows that pronoun ergators do not undergo the rule:

(203) *e mea kino Ø t - a - na patu - nga
ō
a thing bad nom the poss his hit nomin

Ø te wawine.
abs the woman

As in Tongan, this restriction can be explained by the fact that the possessive *a* and the ergative *e* have essentially the same function. Both mark the NP as a semantic agent, and so a rule which replaces one with the other would simply be redundant. On the other hand, the restriction may also be caused by the fact that the ergative is special and relatively new in Pukapukan. This view is somewhat supported by the informant, who insisted that (204) was 'grammatical', but only used by some people:

(204) e mea kino Ø te patu - nga e Turi
a thing bad nom the hit nomin erg Turi

Ø te wawine.
abs the woman

"The killing by Turi [of]
the woman is a bad thing."

(Notice that the object in (204) is preceded by the particle *Ø*.)

Possessive Marking has a rather different effect on the subjects of intransitive nominalizations. These NP's must undergo the rule, but can be placed in either the a or o form of the possessive. The choice of a or o is not determined by grammatical relations, as is shown by (205-6):

(205) e lelei Ø te lele - nga a te kovi lewu.
a good nom the run nomin poss the pers. little

"The running of the child [past] is a good [thing]."

(206) e lelei Ø te lele - nga o te kovi lewu.
a good nom the run nomin poss the pers. little

"The running of the child [non-past] is a good [thing]."

Compare:

(207) na lele Ø te kovi lewu. "The child ran."
tns run nom the pers. little

Instead, the possessive indicates the tense of the nominalization, with a denoting past and o, non-past action. Since tense is often conditioned by the matrix sentence, there are some contexts where one of the two forms is excluded:

(208) auwaa t - a - u wano - nga e mamao.
*o - o
maybe the poss you go nomin tns far

"Perhaps your going (= the amount you will have gone) is far."

Compare:

(209) e kino Ø t - a - u wano - nga ki Aakalana.
o - o
a bad nom the poss you go nomin to Auckland

"Your having gone/going to Auckland is a bad [thing]."

The possessive can also change the focus of the nominalization from the action (o) to the manner in which the action is accomplished (a). Consider (210), taken from a Pukapukan tale (Beagleholes, 1938):

(210) 0 onoono 0 koolua i t - a - ku luku - nga
tns watch tns you du. acc the poss my dive nomin

ki lalo o te moana.
to under of the ocean

"You (du.) watch [the manner of] my diving below the ocean."

Ideally, this semantic difference should be represented in some way at the level of deep structure. But this representation would fail to capture the fact that Possessive Marking in (197) and (205) is essentially the same. The form of the possessive is fixed in (197) because it must indicate the grammatical relation of the noun to the verb (a = actor, o = object). It is only because these relations are neutralized for intransitives that a semantic choice is allowed in (205).

These facts suggest that all instances of Possessive Marking should be explained by a single set of transformations. The change of meaning which is sometimes introduced will be accounted for by an interpretive rule:[56]

(TP2) Possessive Marking (*a*)

$$[_{NP}\ \underbrace{X\ \underline{(C)(a)nga}}_{1}\ \ \underset{2}{\underline{\emptyset}}\ \ \underset{3}{NP}\ \ \underbrace{(\underline{i}\ NP)}_{4}\ \ \underbrace{Y}_{5}\ {}_{NP}]$$

$$\longrightarrow\ 1\ \ \underline{a}\ \ 3\ \ 4\ \ 5$$

oblig. if 4 ≠ Ø
opt. if 4 = Ø, and
5 ≠ *e* NP

Possessive Marking (*a*) is required for all actors of accusative verbs, and allowed for subjects of intransitives. Then, Possessive Marking (*o*) places other NP's closest to the verb in the *o* form of the possessive:

(TP3) Possessive Marking (*o*)

$$[_{NP}\ \underbrace{X\ \underline{(C)(a)nga}}_{1}\ \ \underset{2}{\begin{Bmatrix}\underline{\emptyset}\\ \underline{i}\end{Bmatrix}}\ \ \underbrace{NP\ \ X}_{3}\ {}_{NP}]\ \longrightarrow\ 1\ \ \underline{o}\ \ 3$$

Notice that neither of these rules applies to ergative nominalizations, and Possessive Marking (*o*) is optional for NP's preceded by an overt *i*. This particular construction is discussed below.

If a nominalization is not passive, its actor may be postposed before either of the Possessive Marking rules has applied. Let us postpone for a moment the details of this Extraposition rule, which produces examples like the following:[57]

(211) e mea kino Ø te patu - nga *i* te wawine
a thing bad nom the hit nomin acc the woman

e Turi.
agt Turi

"The killing [of] the woman
by Turi is a bad thing."

Possessive Marking may then apply, and optionally places the object in the o form of the possessive. The nominalization which results from this has parallels in Tongan and Maori:

(212) e mea kino Ø te patu - nga o te
*a
a thing bad nom the hit nomin poss the

wawine e Turi.
woman agt Turi

"The killing of the woman by Turi is a bad thing."

As in Maori, (212) differs from a passive nominalization (198) because it does not have a passive suffix, and need not undergo the Possessive Marking rule. (Possessive Marking is required for objects of passive nominalizations, which are preceded by the particle Ø.) (213) shows that the object of a passive cannot take the particle i instead:

(213) *e mea kino Ø te patu - a - nga i te
a thing bad nom the hit pass nomin acc the

wawine e Turi.
woman agt Turi

The nominalization in (212) also has an alternative form, in which the actor rather than the object appears in the possessive (197). But passive nominalizations are not allowed this option:

(214) *e mea kino Ø te patu - a - nga a Turi
o
a thing bad nom the hit pass nomin poss Turi

i te wawine.
Ø
o
the woman

Finally, the two types of nominalizations have slightly different meanings, since (198) implies duration but (212) does not. The nominalization in (198) means that Turi hit the woman several times; that in (212) means that he hit the woman only once.

Since (198) and (212) have such different characteristics, they cannot both have been derived by means of the passive rule. Instead, (211-2) must have undergone a separate rule of Extraposition, which operates in one of two ways. Either the rule applies to an accusative nominalization, by reordering actor and object and marking the actor with e. This is essentially the way that Extraposition operates in Maori:

(TP4a) Agent Extraposition I (optional)

$$[_{NP}\ \underbrace{X\ \underline{(C)(a)nga}}_{1}\quad \underbrace{\underline{\emptyset}}_{2}\quad \underbrace{NP}_{3}\quad \underbrace{\underline{i}\ NP}_{4}\quad \underbrace{Y_{NP}}_{5}]$$

$$\longrightarrow 1\quad 4\quad \underline{e}\quad 3\quad 5$$

Or the rule might possibly apply to an ergative nominalization, by reordering actor and object and marking the object with i. This process is something like Extraposition in Tongan:

(TP4b) Agent Extraposition II (optional)

$$[_{NP}\ \underbrace{X\ \underline{(C)(a)nga}}_{1}\quad \underbrace{\underline{e}\ NP}_{2}\quad \underbrace{\underline{\emptyset}}_{3}\quad \underbrace{NP}_{4}\quad \underbrace{Y\ _{NP}}_{5}]$$

$$\longrightarrow 1\quad \underline{i}\quad 4\quad 2\quad 5$$

It is interesting that these two versions of Extraposition require exactly the same amount of formal machinery. Moreover, there is almost no evidence in Pukapukan which supports one version over the other. One might claim that (211) should not be derived by (TP4b), since it is distinctly more acceptable than an ergative nominalization. On the other hand, the nominalization required by (TP4a) is also ungrammatical, unless its actor has been placed in the possessive.

We are left with the peculiar fact that (211-2) can be generated in either of two adequate ways. This conclusion may seem unsatisfactory for a description of Pukapukan, but it is particularly suggestive from a comparative viewpoint. We have found nominalizations like (212) in both accusative and ergative languages, although these languages share no other types of nominalizations. This distribution suggests that (212) could survive the accusative-to-ergative drift exactly because it could be derived in either of two ways. In accusative languages, (212) would be derived by (TP4a); it would be reanalyzed as (something like) (TP4b) in languages which had generalized the ergative.

Finally, a rule of Possessive Preposing applies after all of the other nominalization rules. This rule is optional, and moves possessive pronouns to a position between the tense marker and the nominalized verb:

(215) e mea kino Ø t - a - na patu - nga
*o
a thing bad nom the poss his hit nomin

i te wawine.
acc the woman

"His killing the woman is
a bad thing."

(216) e mea kino 0 t - o - na patu - a - nga
*a
a thing bad nom the poss his hit pass nomin

e Turi.

"His being killed by Turi
is a bad thing."

(217) e mea kino 0 t - o - na patu - nga
*a
a thing bad nom the poss his hit nomin

e Turi.
agt Turi

"His killing by Turi is
a bad thing."

Examples of the rule applying to intransitives are given in (209). Since Preposing does not affect the form of the possessive, it can be formalized quite simply:

(TP5) Possessive Preposing (optional)

$$\underbrace{[_{NP}\ \text{Art}}_{1}\quad \underbrace{\text{V}\ \underline{(\text{C})(\text{a})\text{nga}}}_{2}\quad \underbrace{\text{Poss [+pro]}}_{3}\quad \underbrace{\text{X}\ _{NP}]}_{4}$$

$$\rightarrow 1\ \ 3\ \ 2\ \ 4$$

The discussion has shown that Pukapukan has four nominalization rules, which are listed below in the order in which they apply:

Nominalization Formation

Agent Extraposition (I or II) (optional)

Possessive Marking (a)

Possessive Marking (o)

Possessive Preposing (optional)

These differ from the Maori rules in three important ways: (1) They allow for ergative nominalizations as well as passive or accusative ones (this is not surprising, since Pukapukan has all three types of case marking). (2) The extraposed nominalizations can be derived from the ergative or accusative types. (3) Possessive Marking has become a semantic rule, since it can introduce a difference of meaning for subjects of intransitive nominalizations. This last is clearly a product of the fact that Possessive Marking is neither ergative nor accusative in Pukapukan. It can identify the subject of an intransitive with either an actor or an object, and thus is very different from Maori or Tongan.

4.4 Samoan. The nominalizations in Samoan deviate the most from the Maori type. Objects of transitive verbs are still limited to the o form of the possessive, but subjects of intransitives, and some actors, may take either a or o. Objects may appear in the possessive when they are not next to the nominalized verb, so that a transitive nominalization is often modified by two possessors. As if to compensate for this syntactic freedom, some important restrictions are placed on the Preposing rule.

The nominalizations are formed by a rule of Nominalization Formation, which replaces the tense marker of the complement with the definite article le. A nominalizing suffix is no longer attached to the verb, but its appearance in lexicalized nouns shows that it was once productive:

(TS1) Nominalization Formation

$$[_{NP}\ [_{S}\ \text{Tns}\quad X\ _{S}]\ _{NP}] \longrightarrow 1\ \ \underline{\text{le}}\ \ 4\ \ 6$$

1 2 3 4 5 6

The lexicalized noun in (218) is formed with the suffix -(C)(a)nga:

(218) po 'o fea 'o iai 0 l - $\frac{o}{*a}$ - na nofo - anga.
Q tns where tns be nom the poss his stay nom.
"Where is his place-at-the-table?"

Compare (219), where the nominalization is derived in the usual way:

(219) 'ua leva ai 0 l - $\frac{o}{a}$ - u nofo 'i Saava'i.
tns long rel nom the poss you stay in Saav.
"Has your staying in Saava'i been long?"

Although the rule of Nominalization Formation does not introduce a suffix, it is interesting that transitive nominalizations are preferred with the passive -(C)(i)a. This preference is much stronger than in simple sentences, where -(C)(i)a may be added or omitted rather freely:[58]

(220) 'ua ou saauni 'i l - a - na sele - ina
tns I prepare acc the poss his cut pass
o lo'u ulu.
poss my hair
"I was prepared for his cutting of my hair."

Compare the simple sentence:

(221) 'ua na sele -(*ina*) Ø lo'u ulu. "He cut
tns he cut pass nom my hair my hair."

Moreover, -*ina* seems to nominalize the verb in (220) instead of placing it in the passive. This fact is established by my informants, who translated the isolated passives in (222) as nominalizations:

(222)	vala'au-ina	'the calling'
	vala'au-lia	'the inviting'
Cp:	vala'au	'call (trans.)'

Examples like this can be best explained by assuming that -*(C)(i)a* has been generalized to any rule which affects the syntactic properties of the verb. This extension of -*(C)(i)a* is syntactically restricted, since the suffix may not be attached to intransitive nominalizations (cf. §1.4). Nonetheless, it does suggest that the nominalizing and passive suffixes are felt to be functionally the same. We have seen evidence for this parallel in languages besides Samoan.

If the NP closest to the verb is a pronoun, it can be moved to second position in the nominalization by a rule of Pronoun Preposing. This rule differs from Possessive Preposing in Maori in preceding the rule which turns NP's into possessors. Thus the clitic in (223) is not in the possessive, but is probably still in the absolutive case:

(223) pe 'e te fiafia 'ona 'o le Ø maatou
Q you tns happy because pred the nom we

alofa 'i le teine.
love acc the girl

"Are you happy because of
our loving the girl?"

Compare:

(224) 'ua maatou alofa 'i le teine. "We love
tns we love acc the girl the girl."

The absolutive (or nominative) seems to be the case of all clitics in Samoan, which are never preceded by the ergative *e*. (225) shows that this is true even for actors of transitive verbs:

(225) e lelei Ø le *Ø* maatou kuka -ina o le i'a.
*e
tns good nom the we cook pass poss the f.

"Our cooking of the fish
is good."

Compare:

(226) 'ua kuka -ina *e* maatou Ø le i'a.
tns cook pass erg we abs the fish

"We cooked the fish."

Since clitics undergo Possessive Marking like the subjects of intransitive verbs, it is easiest to assume that they are preceded by *Ø*, rather than by no particle at all. This allows us to formalize Pronoun Preposing as follows:

(TS2) Pronoun Preposing (optional)

$[_{NP}$ Art V Case [+pro] X $_{NP}]$

1 2 3 4 5

$\longrightarrow$ 1 $\underline{\emptyset}$ + $\overset{4}{[\text{+clit}]}$ 2 5

It is important that this rule applies to pronouns which are actors (of transitives) or subjects (of intransitives), but never to objects of transitive verbs. The fact that objects are not preposed is demonstrated by (227):

(227) *'ua ou saauni 'i l - o - 'u 'otengia - ina
a
0
tns I prepare acc the my reprove pass

e le faife'au.
agt the minister

Compare (228), where the rule has not applied:

(228) 'ua ou saauni 'i le 'otengia - ina o a'u
*a
tns I prepare acc the reprove pass poss me

e le faife'au.
agt the minister

"I am prepared for the scolding of me by the minister."

From:

(229) 'ua 'otengia-ina 0 a'u e le faife'au.
tns reprove pass nom I agt the minister

"I was scolded by the minister."

On a formal level, this fact can be explained by ordering Pronoun Preposing before the Scrambling and Extraposition rules. This ordering is apparently more crucial than that of any other nominalization rules in Samoan. This is established by the informants, who claimed that (228) was completely meaningless and could not be corrected at all. It is also suggested by the relation of Preposing to other rules which are discussed below.

After Preposing has applied, the NP's modifying the verb may be placed in the possessive. Possessive Marking is optional for objects of transitive verbs, as shown by (230) and (232):

(230) e lelei Ø le aumai -(ina) e Ioane Ø le tusi.
tns good nom the bring pass agt John nom the book

"The delivering by John
[of] the book is good."

From:

(231) 'ua aumai -(ina) e Ioane Ø le tusi.
tns bring pass agt John nom the book

"John delivered the book."

(232) 'ua leva Ø le lee va'ai a le fooma'i
o
tns long nom the not see poss the doctor

'i a maatou.
acc pers us

"It's been long, the doctor's not seeing us (=
it's been long since..)"

From:

(233) e lee va'ai Ø le fooma'i 'i a maatou.
tns not see nom the doctor acc pers us

"The doctor doesn't see
us."

(232) resembles a transitive nominalization in Maori; (230) resembles a transitive nominalization in Tongan.

If Possessive Marking does apply to the object, it is always placed in the *o* construction. The other form of the possessive is not allowed, as is illustrated by (234-5):

(234) e lelei Ø le aumai -(ina) e Ioane o le tusi.
*a
tns good nom the bring pass agt John poss the b.

"The delivering by John
of the book is good."

(235) 'ua leva Ø le lee va'ai a le fooma'i
o
tns long nom the not see poss the doctor

o maatou.
*a
poss us

"It's been long, the doctor's not seeing of us."

This part of Possessive Marking is thus a syntactic rule which resembles Possessive Marking in Maori. However, it differs from the Maori rule in not requiring that the possessor be the NP closest to the nominalized verb. As we have seen, an object possessor may be separated by the verb by an ergative or nominative NP.[59] Since this last type of noun can also undergo the rule, it is normal for nominalizations in Samoan to have more than one possessor.

Possessive Marking can also apply to the actors of passive or ergative verbs, although this construction is not particularly common. The actor must be the NP closest to the nominalized verb, and is placed in the a construction:

(236) e lelei Ø le kuka -(ina) a Sali o le i'a.
*o
tns good nom the cook pass poss Sali poss the f.

"The cooking of Sali of the fish is good."

From:

(237) na kuka -(ina) e Sali Ø le i'a.
tns cook pass agt Sali nom the fish

"Sali cooked the fish."

Since sentences like (236) are not allowed in Tongan, it is interesting that they seem restricted in Samoan as well.

My informants accepted (236), but claimed that (234) was much better. Constructions like (236) are also restricted on a grammatical level, since they require that the object NP also be in the possessive. (238) shows what happens if the object remains in the absolutive (or nominative):

(238) *e lelei Ø le aumai -(ina) a Ioane
ó
tns good nom the bring pass poss John

Ø le tusi.
nom the book

The reasons behind this specific restriction are unclear to me. However, it does support the observation that ergative nouns are not usually placed in the possessive in Polynesian (v. §4.2).

Finally, Possessive Marking must affect the actors of accusative verbs, the subjects of intransitives and any preposed actor or subject. (Actually, the rule is optional for preposed pronouns which are non-singular, but obligatory in all other cases.[60]) These NP's are the first nouns in their nominalization and are preceded by Ø, so they undergo Possessive Marking in a uniform way. They may be placed in the o possessive, as in (239):

(239) 'ua ou ofo 'i le alofa o le tama
tns I surpr. acc the love poss the boy

'i le teine.
acc the girl

"I was surprised at the boy's
loving the girl (= the
fact that he loved her)."

From:

(240) saa alofa Ø le tama 'i le teine.
tns love nom the boy acc the girl

"The boy loved the girl."

Or, they may be placed in the a possessive, as in (241):

(241) 'ua ou ofo 'i le alofa a le tama 'i
tns I surpr. acc the love poss the boy acc

le teine.
the girl

"I was surprised at the boy's
loving the girl (= the
way he loved her)."

As in ordinary possessive constructions or Pukapukan nominalizations, the choice of a or o introduces a change in meaning. A implies that the possessor is dominant or active towards the verb; o implies that he is more neutral. Thus a may indicate that the nominalization is past tense, as opposed to o, which does not:

(242) e lee se mea leanga Ø l - a - na oso
tns not a thing bad nom the poss her jump

i tua.
in back

"It could have been a bad
thing if she hadn't jumped
in back."

But:

(243) e lee se mea leanga Ø l - o - na oso
tns not a thing bad nom the poss her jump

i tua.
in back

"Her jumping in back is
not a bad thing."

Compare the simple sentence:

(244) 'ua na oso i tua. "She jumped in
tns she jump in back back."

A can also focus on the manner of an action, while o just indicates that it occurs:[61]

(245) e mea leanga Ø le ulaula-tapa'a o
a thing bad nom the smoke tobacco poss

le faife'au.
the minister

"The fact that the minister smokes tobacco is a bad thing."

But:

(246) e maatangaa Ø le ulaula-tapa'a a le
tns ugly nom the smoke tobacco poss the

faife'au.
minister

"The way the minister holds his cigarette is ugly."

It is thus not surprising that certain semantic contexts allow only one of the two types of possessive. For instance, o was rejected in (248) by some of my informants, who claimed that 'everyone knows how to read':

(247) 'ua laatou ofo 'i l - a - na faitau-tusi.
tns they surpr. acc the poss him read book

"They were surprised at the way he reads books."

But:

(248) ?*'ua laatou ofo 'i l - o - na faitau-tusi.
tns they surpr. acc the poss him read book

"They were surprised at the fact that he reads books."

But o is preferred for the actors of most accusative verbs, which describe emotions or perceptions:

(249) saa ou maafaia Ø le mea lenaa 'ona 'o
tns I manage nom the thing that bec. pred

l - o - 'u mana'o iai.
the poss my want rel

"I managed that thing because of the fact that I wanted to."

But:

(250) ?*saa ou maafaia Ø le mea lenaa 'ona 'o
tns I manage nom the thng that bec. pred

l - a - 'u mana'o iai.
the poss my want rel

"I managed that thing because of the way that I wanted to."

Finally, the semantic contrast between a and o is usually neutralized for the preposed actors of ergative or passive verbs. This is because all passive/ergative verbs must take a semantic agent, which is by definition active and agentive. My informants volunteered the different translations of (251-2), but indicated that for other such verbs, an a or o would "make no difference:"

(251) e lelei l - o - na kuka -(ina) o le i'a.
tns good the poss his cook pass poss the fish

"It's good if he cooks the fish."

And:

(252) e lelei l - a - na kuka -(ina) o le i'a.
tns good the poss his cook pass poss the fish

"The way he cooks the fish is good."

From:

(253) ʻua na kuka -(ina) Ø le iʻa. "He cooked the fish."
tns he cook pass nom the fish

Apparently, the choice of possessive here is less meaningful than for the subjects of intransitive verbs.

It is nonetheless important that both forms of the possessive are allowed, since ergators which are not preposed may only occur in the a possessive (236). I have suggested that this difference can be accounted for on a formal level by having the Preposing rule place the clitics in the Ø case. (Recall that this is the case of all NP's for which the a/o contrast is permitted.) While this solution is descriptively adequate, it is obvious that the difference is motivated by the position of the pronouns.[62] Preposed pronouns have an unambiguous grammatical function because the Preposing rule is limited to subjects (of intransitives) and actors (of transitives). Because pronouns in second position are separated from other NP's in this way, no confusion is caused by allowing them to take either (or neither) form of the possessive.[63]

I should add that these considerations always seem to dictate whether a, o, or both forms of the possessive are allowed. For instance, intransitives are only modified by one grammatical NP, so this NP may take a or o regardless of its position. But transitives are subcategorized for two NP's, so these may take both forms of the possessive only when they have been "removed." It is interesting that this possibility is never open to objects, which are always restricted to o. This seems to be because they always follow the nominalized verb, where they can always potentially be confused with an actor.

Despite the complications which it seems to involve, Possessive Marking can be formalized as two reasonably simple rules. The first rule affects the leftmost NP of a nominalization, which it optionally places in the a possessive:

(TS3) Possessive Marking (a) (optional)

$$\underbrace{[_{NP}\ \text{Art}\ \ X}_{1}\ \ \underset{2}{\begin{Bmatrix}\underline{e}\\ \underline{\emptyset}\end{Bmatrix}}\ \ \underbrace{NP\ \ Y\ (\underline{\emptyset}\ NP)\ Z\ _{NP}]}_{3} \longrightarrow 1\ \ \underline{a}\ \ 3$$

(TS3) collapses two disjunctively ordered rules, and thus can never affect the object of a transitive verb. Since the ergator in (236) is also a leftmost NP, it is included in the same rule as preposed subjects and the actors of accusative verbs.[64]

The second rule places NP's preceded by Ø or i in the o form of the possessive. It is optional for plural clitics and the objects of certain verbs, but obligatory elsewhere:

(TS4) Possessive Marking (o)

$$\underbrace{[_{NP}\ \text{Art}}_{1}\ \ \underset{2}{X}\ \ \underset{3}{\begin{Bmatrix}\underline{\emptyset}\\ \underline{i}\end{Bmatrix}}\ \ \underset{4}{NP}\ \ \underbrace{Y\ \ _{NP}]}_{5} \longrightarrow 1\ \ 2\ \ \underline{o}\ \ 4\ \ 5$$

opt. if 3 = i,
or 2 = $\bar{Z}$ e NP,
or 4 = $\begin{bmatrix}+\text{clit}\\ +\text{pl}\end{bmatrix}$
oblig. otherwise

Except for the conditions on clitics, which were mentioned above, the details of (TS4) agree completely with the general

Polynesian restrictions which I have described. (TS4) guarantees that all actors or subjects which are marked with 0 will appear with an overt (possessive) particle in surface structure. The rule is optional for accusative objects, which are preceded by i. It is optional for 0 objects, because 0 in this instance does not mark a true nominative case. (This is true even for objects of passive verbs, since the passive in Samoan is a low-level variant of the ergative.) The fact that these conditions survive, although the rules are very different, shows that Possessive Marking in Samoan is the same kind of process as in Maori.

Finally, a rule of Extraposition may apply before Possessive Marking (o). This rule optionally permutes the actor and object of a transitive nominalization, if the actor is preceded by e and the object, by 0:

(254) e lelei 0 le aumai -(ina) o le tusi
tns good nom the bring pass *a the book
e Pita.
agt Pita

"The delivering of the book by Pita is good."

Then, the object must be placed in the o form of the possessive, as illustrated by (254). The result bears a surface resemblance to constructions in Maori and Pukapukan, but in Samoan the order of Extraposition and Possessive Marking (a) is reversed. Extraposition may not apply if the actor is in the possessive:[65]

(255) *e lelei Ø le aumai -(ina) o le tusi
Ø
tns good nom the bring nomin — the book

a Pîta.
poss Pita

and so Extraposition must follow (TS3).

Extraposition must precede Possessive Marking (o), since sentences like (256) are not allowed:

(256) *e lelei Ø le aumai -(ina) Ø le tusi
tns good nom the bring pass nom the book

e Pita.
agt Pita

Moreover, the rule must not affect accusative nominalizations, as demonstrated by (257):

(257) *'ua ou ofo 'i le alofa o le teine
'i
tns I surpr. acc the love — the girl

e le tama.
Ø
— the boy

These facts allow us to formalize Extraposition as follows:

(TS5) Agent Extraposition (optional)

$$[_{NP}\ \underbrace{\text{Art V}}_{1}\ \ \underbrace{\underline{\text{e}}\ \text{NP}}_{2}\ \ \underbrace{\underline{\emptyset}\ \text{NP}}_{3}\ \ \underbrace{\text{X}}_{4}\ _{NP}] \longrightarrow 1\ \ 3\ \ 2\ \ 4$$

We have seen that the Samoan rules obey several strict ordering relations. Pronoun Preposing must precede the Possessive Marking rules, to account for the differences in possessives between ergative clitics and ergative nouns. Possessive Marking (a) must precede Extraposition, because objects may never take the a possessive, even when they are closest to the nominalized verb (254). These facts allow us to order the rules for nominalizations:

Nominalization Formation
Pronoun Preposing (optional)
Possessive Marking (a) (optional)
Agent Extraposition (optional)
Possessive Marking (o)

These rules differ from the Maori rules in a number of important ways: (1) the rule of Nominalization Formation does not introduce a nominalizing suffix, (2) Pronoun Preposing never affects the object of a transitive verb, (3) the Possessive Marking rules refer to the leftmost NP, rather than the NP closest to the verb,[66] and (4) the rules of Possessive Marking (a), Extraposition and Preposing have the reverse order from their order in Maori. These characteristics make the nominalization process in Samoan considerably different from any other Polynesian language. But the survival of details like the conditions on Possessive Marking show that the Samoan process is the same in spirit.

These characteristics also combine to make Possessive Marking neither ergative nor accusative for most nouns. Full nouns which are the subjects of intransitives may undergo the rule like actors (a) or objects (o) of transitive verbs. The difference in meaning associated with this choice suggests that Possessive Marking is a semantic rule, as in Pukapukan. On the other hand, the rule is accusative for clitic pronouns, since both actor and subject clitics have a choice between a and o. This accusative structure is motivated by the structure of Pronoun Preposing, which applies to pronominal actors or subjects, but not to pronominal objects.

4.5 Summary. The data discussed in this survey of nominalizations is summarized on p. 104, and shows quite clearly that the nominalization process can be reconstructed for proto-Polynesian. Although the details of the process vary from language to language, each has some version of the four basic nominalization rules: Possessive Marking, Nominalization Formation, Extraposition and Possessive Preposing. These rules must thus be descended from one proto-Polynesian model, which looks the most like the nominalization process in Maori.

A much more crucial question from our point of view is how the nominalizations in each language came to deviate from this model. The survey has suggested that most of this change was motivated by the accusative-to-ergative drift.

4.5.1 Possessive Marking. The influence of the drift is most obvious in the changes in Possessive Marking, which places an NP modifying the verb into the possessive. Possessive Marking is accusative in Tongan, an ergative language; it is ergative in Maori, where simple sentences are accusative. In Samoan and Pukapukan, which have a mixed type of case, NOMINALIZATION in

	Maori	Pukapukan	Samoan	Tongan
1.	Nominalization Formation	Nominalization Formation	Nominalization Formation	Nominalization Formation
2.	Agent Extraposition	Agent Extraposition (I, II)	Pronoun Preposing	Subject Possessive Marking
3.	Possessive Marking	Possessive Marking (a)	Possessive Marking (a)	Agent Extraposition
4.	Possessive Preposing	Possessive Marking (o)	Agent Extraposition	Object Possessive Marking
5.		Possessive Preposing	Possessive Marking (o)	Possessive Preposing
6.				Emphatic Copy Deletion

Differences:

NOMINALIZATION FORMATION introduces a nominalizing suffix in Maori and Pukapukan; it does not introduce a suffix in Tongan or Samoan.

AGENT EXTRAPOSITION applies to accusative sentences in Maori, but to ergative (or passive) sentences in Samoan and Tongan. It may apply to either accusative or ergative sentences in Pukapukan.

POSSESSIVE MARKING is a syntactic rule in Maori and Tongan; it introduces a change of meaning in Pukapukan and Samoan. The rule refers to the leftmost NP in nominalizations in Samoan, but to the NP closest to the verb in nominalizations in other languages.

POSSESSIVE PREPOSING is obligatory for pronouns in Tongan, but optional in all other languages. It may also apply to full noun possessors in Maori. The rule is restricted to actors and subjects (not necessarily possessors) in Samoan, but can apply to all grammatical NP's which are possessors, in other languages.

the rule does not identify the subject with either an actor or an object. Instead, both forms of the possessive are allowed, and the choice of one or the other is determined by semantics. (This is exactly how possessives are chosen for ordinary possessive constructions.) The four rules can thus be arranged in a continuum:

	Maori	Puk.	Samoan [-clit]	Samoan [+clit]	Tongan [-clit]	Tongan [+clit]
Actor	a	a	a	a/o	-	a
Object	o	o	o	o	o	o
Subject	o	a/o	a/o	a/o	a	a
Case Struct of Rule	erg	mixed	mixed	acc	-	acc

As this is basically the reverse of the case marking continuum in §1.5, the changes in Possessive Marking must be caused by the accusative-to-ergative drift. This influence is what we might expect, since Possessive Marking does assign the case relations in nominalizations. A more important fact is that the changes are designed so that the rule has a single relation to simple sentence case. This relation is captured in the generalization that

(A) The case structure of Possessive Marking is the opposite of the (normal) case marking of simple sentences.

(A) demonstrates that the case marking rules are systematically connected to one other transformation in Polynesian. As such, it suggests that case marking is non-arbitrary at some level of grammatical structure. It also adds support to my claim that case is derived, since the case structure of Possessive marking is always the opposite of the case of simple sentences. If accusative relations were more basic than ergative ones, we would expect not to find this balance in other (presumably deeper) parts of the grammar.

(A) clearly does not qualify to be a universal, since some languages have a single possessive form for all NP's in a nominalization.[67] On the other hand, the generalization does hold for a number of languages outside of Polynesian. Possessive Marking is ergative for nominalizations in Latin and Russian, but the case marking in these languages is clearly accusative. Examples like these suggest that (A) should be restated as (A'):

> (A') If the Possessive Marking rule has a case structure, it will be the opposite of the (normal) case marking of simple sentences.

4.5.2 Restrictions on Possessive Marking. The restrictions on Possessive Marking are also affected by the accusative-to-ergative drift. We have seen that Possessive Marking is obligatory in Maori for subjects of intransitives, actors of accusative and objects of passive verbs. In other words, the rule obeys a general constraint that all NP's in nominalizations must be preceded by overt markers. This constraint survives in Samoan and Pukapukan for accusative (and intransitive) nominalizations, but does not affect the ergative constructions at all.[68] In Tongan, where accusative sentences have been reanalyzed as intransitive, both subjects and objects of nominalized verbs may be preceded by 0 (177, 181).

Clearly, the elimination of the 0 restriction is related to the change in the case marking of simple sentences. If we allow for some freedom in the connection of the two, we can state the constraint as follows:

> (B) If there is a case marking for simple sentences where 0 classifies logical subjects, then all NP's in nominalizations of this type must be preceded by overt markers.

(B) shows that the case marking of simple sentences can also organize the constraints on other rules. It thus provides more evidence that case in Polynesian is meaningful at some level of grammar.

4.5.3 Possessive Preposing and Extraposition. Finally, Possessive Preposing and Extraposition are affected in similar ways by the accusative-to-ergative drift.

We have seen that Extraposition in Maori bears a formal resemblance to the passive; both rules may optionally extrapose the actor, which is marked with the particle e. Since the rules affect no other NP's besides the actor of a transitive verb, they may be said to have an ergative structure. They are, moreover, radical rules, since they prevent the actor from being placed in its normal nominative (or possessive) case. This functional importance is indicated by the fact that the rules are ordered relatively early in the list of transformations.[69]

However, both rules lose this radical structure in other parts of the continuum. Extraposition in Samoan must be analyzed as just a scrambling rule for ergative clauses; it must also be analyzed that way in Tongan. Since this scrambling rule does not affect the actor any more than the object, it cannot be said to have a case structure at all. Similarly, the passive in Samoan can be analyzed as a suffix attachment rule, which adds -(C)(i)a to any transitive verb. The suffix rule is completely lexicalized in Tongan, where it need not cause a change in case marking, and thus is neither an accusative nor ergative rule.

This loss of case structure is paralleled by a change in the importance of the two rules. For instance, Extraposition comes to have little effect on which NP's can undergo the Possessive Marking rule. In Tongan, actor pronouns are possessivized before Extraposition has even applied; in Samoan, objects may be possessivized regardless of their position in the nominalization. The passive ceases to have an effect on the case marking of transitive clauses, which take the ergative in Tongan even when they do not have a suffix on the verb.

We can guess that the disintegration of these two rules is related to the fact that both have an ergative case structure in accusative languages. The motivating principle behind the change is then

(C) Optional clause restricted rules may not have the same case structure as the case marking of simple sentences.

(C) explains why Extraposition becomes a scrambling rule in Tongan and Samoan, and why the passive is reduced to a rule of suffix attachment in ergative languages. We can add to it the observation that Extraposition is functionally a kind of passive. Like the Passive, Extraposition causes a radical change in the case marking of clauses in an accusative language. It is an interesting question whether the ergative languages in Polynesian can also have a radical rule.

Possessive Preposing is also affected by the drift. We have seen that the rule applies to all types of possessors in Pukapukan and Maori, and thus does not have a particular case structure. Preposing in Tongan, on the other hand, cannot apply to pronominal objects if pronominal actors can undergo the rule. (This is a direct consequence of the fact that Subject Possessive Marking precedes Extraposition, and is obligatory for pronouns.) This accusative structure is even clearer in Samoan, where Preposing is limited to pronouns which are subjects or actors (but not objects). Apparently, Preposing is accusative in ergative languages, just as Extraposition is ergative in accusative ones.

It is important that this change of case seems to be accompanied by a change in the importance of the rule. Preposing is relatively low-level in Pukapukan and Maori, where it applies after all other nominalization rules. But

Preposing in Samoan is ordered relatively early, before the Possessive Marking rules. Moreover, it determines the types of possessive in which a pronominal actor may be placed (§4.4).

We can conclude, therefore, that Possessive Preposing also obeys restriction (C); it is accusative in basically ergative languages, but has no case structure in accusative ones. The fact that this distribution is related to functional importance suggests that Preposing is a kind of antipassive. Like Extraposition or the Passive, Preposing is a radical rule in languages which have the opposite type of case marking (e.g. Samoan). Although it differs from Extraposition in being restricted to pronouns, the motivating principle behind the two rules is essentially the same.

4.5.4 It thus appears that most of the variation in nominalizations is caused by the accusative-to-ergative drift.[70] Possessive Marking is constrained so that its case structure will always be the opposite of that of simple sentences. Extraposition and Possessive Preposing are also constrained by restrictions which refer to the case marking of simple sentences. Since these restrictions apply equally in accusative and ergative languages, neither type of case marking can be said to be more basic than the other. Instead, both types must be derived at the same (or a similar) level of intermediate structure.

Constraints (A-C) also establish that case marking cannot be completely arbitrary, since it is closely linked to the case structures of other transformational rules. Case marking gives the opposite paradigm from that of Possessive Marking; it never gives the same paradigm as that of Preposing or Extraposition. Consequently, all four rules operate in a functionally coherent way which is determined by the organizing principle of case. It is an interesting but unsolved question why Polynesian does not allow the same paradigm for case marking and any other clause restricted rules.

4.5.5 Finally, we have seen that the nominalization rules are often similar to rules for simple sentences. Possessive Marking is a kind of case marking rule for nominalizations; Extraposition acts as a reduced type of passive. These parallels are not terribly surprising when we consider that nominalizations are clauses in underlying structure; however, they do suggest that the generalization could be extended in an important way. If Preposing, like Extraposition, has a parallel in simple sentence transformations, this rule might turn out to be a kind of anti-passive.

In the next section I try to show that Clitic Placement in ergative languages is something of this sort of rule. Clitic Placement does, first of all, resemble Possessive Preposing, since it uses the same morphological set of pronouns and is accusative in ergative languages. (The parallel is especially clear in Tongan, where both rules are allowed to leave emphatic copies behind.) The evidence also suggests that Clitic Placement functions as an anti-passive, although there are several crucial differences which separate the two rules. These differences are discussed in detail below.

5.0 Clitic Placement

Three of the four representative languages have a rule of clitic placement, which moves a pronoun to the second position in its clause. This statistic is slightly misleading from a typological point of view, since the rule has been lost in most other Polynesian languages.[71] Nonetheless, clitic placement can be reconstructed for proto-Polynesian, and its structure in particular languages is affected by the accusative-to-ergative drift.

5.0.1 One fact about clitic placement which suggests its age and original distribution is that it uses a distinct morphological set of pronouns. These clitics are used in special, preposed positions in the sentence, and can probably be attributed to proto-Oceanic. A few members of the set survive in particular Polynesian languages, but most have been replaced by full pronouns (to which they were formerly opposed).[72] The contrast is reasonably well preserved in Samoan:

Full Pronouns:	sg.	du.	pl.
1 excl.	a'u	('i) maa'ua	('i) maatou
1 incl.		('i) taa'ua	('i) taatou
2	'oe	'oulua	'outou
3	ia	('i) laa'ua	('i) laatou

Clitic Pronouns:	sg.	du.	pl.
1 excl.	'ou; 'u[1]	maa[2]	-[3]
1 incl.		taa[2]	-[3]
2	'e; u[1]	lua[2]	tou[2]
3	na[2]	laa[2]	-[3]

---

1. Used for possessive pronouns.
2. Full pronoun may also be used.
3. Full pronoun must be used.

The chart shows that full pronouns may be used in many cases where clitics are required. They may replace clitics in the dual, and must replace them in the 1st and 3rd plural.[73] The singular clitics are apparently the last to disappear; in this respect, Samoan resembles most other Polynesian languages.

5.1 Maori. Maori and all other Eastern languages lack a rule of Clitic Placement. Consider the following examples, all of which are ungrammatical:

(290) *ka na patu i te taane. ("He hit the man.")
tns he hit acc the man

(291) *i ku patu *Ø* te taane. ("The man hit me.")
tns I hit nom the man

(292) *ka maatou patu - *a* *Ø* te wahine.
tns we hit pass nom the woman
("The woman was hit by us.")

(293) *ka na patu - *a* *e* maatou. ("She was hit by us.")
tns she hit pass agt us

(294) *kua ku haere. ("I have gone.")
tns I go

Despite the fact that Clitic Placement is not a synchronic rule, Maori does have a construction which seems to be descended from it. This construction preposes the actor of a transitive sentence and places it in the *a* possessive. Then, Accusative Marking applies and places the object in the nominative case:

(295) n - *aa* te tangata i patu *Ø* te poaka.
tns poss the man tns hit nom the pig
"*The man* hit the pig."

Notice that only the *a* form of the possessive is allowed:

(296) *n - *oo* te tangata i patu *Ø* te poaka.
tns poss the man tns hit nom the pig

And the object may not be placed in the accusative:

(297) *n - *aa* te tangata i patu *i* te poaka.
tns poss the man tns hit acc the pig

The *naa*-construction emphasizes that the actor is an active, direct cause of the verb (and thus has been called "agentive" by Hohepa, 1967). It operates only on the actors of active transitive sentences. (298) shows that an object may not be preposed instead:

(298) *n - aa te poaka i patu Ø te tangata.
oo e
tns poss the pig tns hit nom the man

Further, the rule may not apply to intransitives:

(299) *n - aa te tangata i haere.
tns poss the man tns go

Nor is it allowed for passive sentences at all:

(300) *n - aa te tangata i patu - a Ø te poaka.
tns poss the man tns hit pass nom the pig

(301) *n - aa te poaka i patu - a Ø te tangata.
tns poss the pig tns hit pass nom the man

The possibility that this construction is descended from Clitic Placement is suggested by its distribution in Polynesian. All Eastern languages, which use the past tense marker i, have some version of the naa-construction but no Clitic Placement rule.[74] On the other hand, no other Polynesian languages seem to have the naa-construction. As the past tense marker in these other languages is na, we can imagine that the naa-construction might have emerged in the following way:

I. na a [+clit] V Ø NP (Nb. the person marker a appears in clitic sent. in Pukapukan)
tns pers

II. n - aa [+clit] i V Ø NP
tns poss tns

III. n - aa NP i V Ø NP
tns poss tns

The diagram suggests that clitic sentences like I. were reanalyzed as possessive sentences like II.; subsequently, the possessive construction was extended from clitics to all actor NP's. The reanalysis could happen because the possessive and personal *a*'s are homophonous; moreover, both clitic (I.) and possessive (II.) constructions use the clitic forms of the pronoun:

(302) n - aa *ku* patu Ø te wahine.
*au
tns poss I hit nom the woman
"*I* hit the woman."

Compare the Pukapukan construction:

(303) na a *ku* kite - a Ø te taane.
*au
tns pers I see pass nom the man
"I recognized the man."

More crucially, the reanalysis is motivated by the fact that *na* was replaced by *i* as a tense marker in East Polynesian. We can imagine that this replacement was straightforward in most cases, but that the clitic sentences were reanalyzed because of the special form of the pronoun. Once the change had accounted for *na* in this way, it added *i* to satisfy the requirement that all verbs be preceded by a normal, overt tense marker.[75]

This version of the reanalysis is admittedly speculative, but it is undeniable that Clitic Placement and the *naa*-construction are historically related in some way. In this respect, it is worth noting that the *naa*-construction has an ergative structure in Maori. The construction preposes actors of transitive verbs, but is allowed to apply to no grammatical NP. We will see that Clitic Placement in Pukapukan operates in exactly this way.

5.2 Tongan. Clitic Placement is required in Tongan for pronouns which are actors of transitive or subjects of intransitive verbs. These pronouns are preposed to second position in the sentence, and are not marked with an overt case particle:

(304) na'a nau oo ki tahi. "They went to the ocean."
tns they go to ocean

(304) shows Clitic Placement applying to the subject of an intransitive. In (305), it moves the actor of a transitive verb:

(305) na'a ku 'ave Ø e tohī. "I took the book."
tns I take abs the book

Notice that neither sentence is acceptable if the rule has not applied:

(306) *na'a oo Ø kinautolu ki tahi.
tns go abs they to ocean

(307) *na'a 'ave 'e au Ø e tohī.
tns take erg I abs the book

Although Clitic Placement is required in these cases, it is not allowed for objects of transitives at all. (308) is ungrammatical because its object has been preposed:

(308) *na'a nau tali 'e he kakaī.
tns they accept erg the people

Compare (309), where the rule has not applied:

(309) na'e tali Ø kinautolu 'e he kakaī.
tns accept abs them erg the people

"The people accepted them."

Clitic Placement therefore resembles the Possessive Preposing rule, which favors actors of transitive nominalizations over objects (§4.2).[76] The resemblance is further supported by the fact that both constructions may leave emphatic copies behind. These copies take the full form of the pronoun, and stress that the pronoun is involved in the action:

(310) na'a nau oo Ø kinautolu ki tahi.
tns they go abs they to ocean

"They themselves went to the ocean."

(311) na'a ku 'ave 'e au Ø e tohī.
TNS I take erg I abs the book

"I myself took the book."

Notice that similar copies are not allowed for objects:

(312) *na'a nau tali Ø kinautolu 'e he kakaī.
tns they accept abs they erg the people

Sentences (310-11) suggest that Clitic Placement in Tongan should be formalized as a series of two rules. The first rule copies the pronoun in the second position of the clause:

(TT7) Clitic Copying (obligatory)

$$\begin{array}{cccc} \text{Tns} & \text{V} & \text{Case} & [\text{+pro}] \\ 1 & 2 & 3 & 4 \end{array} \longrightarrow 1 \overset{4}{[\text{+clit}]}\ 2\ 3\ \overset{4}{[\text{-clit}]}$$

Then, the pronoun which is left behind may be deleted:

(TT8) Emphatic Copy Deletion (optional)

$$\underset{1}{\text{V}}\ \underbrace{\text{Case}\ [+\text{pro}]}_{2} \longrightarrow 1\ \emptyset$$

These rules can obviously be combined with the Possessive Preposing rules, and have two other aspects which are particularly worth noting. First, Emphatic Copy Deletion should normally apply, even though it is only an optional rule. (This preference for deletion affects possessive pronouns as well as clitics in simple sentences.) Second, the Clitic Copying rule has an accusative structure, since it treats the actors of transitives like the subjects of intransitive verbs. Both types of pronouns may undergo the rule, as opposed to pronominal objects, which may not.

5.3 Pukapukan. The previous sections have suggested that the case structure of Clitic Placement may also be affected by the accusative-to-ergative drift. Clitic Placement is accusative in an ergative language (Tongan); from what we can reconstruct, it may have been ergative in an accusative one (Maori). These facts suggest that languages in the middle of the continuum might have yet a third type of Clitic Placement rule. This guess is confirmed by evidence from Pukapukan, where the rule is ergative in a peculiarly restricted way.

Clitic Placement is optional in Pukapukan for pronouns which are actors of passive verbs. These pronouns are preposed and preceded by a, which is the marker of pronouns and proper names:

(313) na a na penapena - ina 0 naa popoa.
tns pers he prepare pass nom the-pl food
"He prepared the food."

This construction seems to be a way of de-emphasizing the actor, since his active participation is stressed if the rule has not applied. Compare (313) with the following:

(314) na penapena - ina e na 0 naa popoa.
tns prepare pass agt he nom the-pl food
"He prepared the food."

It is interesting that actors of passive verbs are apparently the only NP's which can undergo the rule. Clitic Placement cannot affect the object of a passive:

(315) *na a laatou penapena - ina e Turi.
tns pers they prepare pass agt Turi

Nor can it affect the subject of an intransitive verb:

(316) *na a ku lele.
tns pers I run

(317) na lele 0 au. "I ran."
tns run nom I

Accusative sentences do not undergo the rule at all:

(318) *na a na tutuli i te tama.
tns pers he send acc the child

(319) *na a na tutuli 0 te wawine.
tns pers he send nom the woman

But:

(320) na tutuli 0 i a ana i te tama.
tns send nom pers she acc the child
"She sent the child."

And ergative sentences are similarly restricted:

(321) *na a u patu Ø te kulii.
tns pers you hit abs the dog

(322) *na a na patu e koe.
tns pers he hit erg you

Compare:

(323) na patu e koe Ø te kulii. "You hit the dog."
tns hit erg you abs the dog

(321) is particularly interesting, since it provides more evidence that ergatives cannot be derived from passives in Pukapukan. If the difference between (321) and (313) were captured in a low-level deletion rule, we would not be able to motivate why (313) is grammatical, but (321) is not.

As discussed in §2.4, the rule for this paradigm can be formalized in one of two ways. If we decide that the difference between (321) and (313) is crucial, then the SD of Clitic Placement should contain a passive suffix:

(TP7) Clitic Placement I (optional)[78]

$$\begin{array}{lllll} \text{Tns} & \underbrace{\text{V (}\underline{\text{C}}\text{)(}\underline{\text{i}}\text{)}\underline{\text{a}}} & \underline{\text{e}} & \text{[+pro]} & \longrightarrow 1\ \ \underline{\text{a}}\ \ 4\ \ 2 \\ 1 & \quad 2 & 3 & \quad 4 & \end{array}$$

If, on the other hand, the restricted use of the ergative is responsible for (321), the rule should be formalized as follows:

(TP8) Clitic Placement II (optional)

$$\begin{array}{lllll} \text{Tns} & \text{V} & \underline{\text{e}} & \text{[+pro]} & \longrightarrow 1\ \ \underline{\text{a}}\ \ 4\ \ 2 \\ 1 & 2 & 3 & \quad 4 & \end{array}$$

Neither of these analyses alters the fact that Clitic

Placement has an ergative structure in Pukapukan. The rule is restricted to actors of one type of transitive verb, and cannot apply to objects or subjects at all. Moreover, it is ergative with respect to the accusative case marking, since it is limited to passive sentences, which are derived from an accusative type. (Recall that the rule does not apply to ergative constructions at all.) This fact suggests that restriction (C) should be modified as follows:

(C') An optional, clause-restricted rule may never[77] have the same case structure as the basic case marking of the sentences to which it applies.

where basic case marking refers to ergative or accusative types.

5.4 Samoan. Clitic Placement is optional in Samoan, but otherwise much like the Tongan rule. The rule moves pronoun actors or subjects to the second position in the sentence, as illustrated below:

(324) 'ua ou fatu -(ina) Ø le pese.
tns I compose pass nom the song

"I composed the song."

From:

(325) 'ua fatu -(ina) e au Ø le pese.
tns compose pass agt I nom the song

"I composed the song."

(324) shows Clitic Placement applying to the actor of a passive/ergative verb; (326) shows it applying to the actor of an accusative:

(326) saa laatou angaleanga 'i le teine.
tns they ill-treat acc the girl

"They treated the girl badly."

From:

(327) saa angaleanga 0 laatou 'i le teine.
tns ill-treat nom they acc the girl

"They treated the girl badly."

In (328), the subject of an intransitive has undergone the rule:

(328) 'ua na sau. "He came."
tns he come

Compare:

(329) 'ua sau 0 'oia. "He came."
tns come nom he

Notice that these clitics differ from the Pukapukan clitics in not being preceded by a personal marker a. This difference can be attributed to the fact that Samoan, like Tongan, never uses a before absolutive pronouns.

As in Pukapukan, Clitic Placement serves to deemphasize the preposed actor or subject. This effect is most obvious in passive or ergative sentences, where a postposed pronoun is clearly more emphatic. Moreover, the rule never applies to object pronouns, as demonstrated by (330-1):

(330) *'ua laatou fasi -(ina) e maatou.
tns they hit pass agt us

("They were hit by us.")

(331) *saa na angaleanga (iai) 0 laatou.
tns she ill-treat rel nom them

("They mistreated her.")

These facts suggest that Clitic Placement is exactly the same sort of rule as the pronoun Preposing discussed in §4.4. Since there is no obvious reason why these rules should be separated in the grammar, conditions of simplicity dictate that they should probably be collapsed. But this collapsing will cause some radical changes in our view of how Clitic Placement works. In §4.4, I claimed that Preposing could be stated most adequately by allowing it to place the clitics in the absolutive case. This operation helped to account for the fact (which is difficult to account for otherwise) that preposed pronouns are possessivized exactly like subjects of intransitives. But if Preposing is to be combined with Clitic Placement, then the latter should also place its clitics in the $\underline{\emptyset}$ case. This would mean that ergative clitics would be placed in the absolutive case, and consequently that Clitic Placement would be a kind of anti-passive.

Perhaps the strongest argument for this analysis is that otherwise, the grammar would have to contain two practically identical rules. I therefore assume that Clitic Placement and Pronoun Preposing are a single rule:

(TS7) Clitic Placement/Pronoun Preposing (optional)

$$\begin{array}{cccccccc} \left\{\begin{matrix}\text{Tns}\\ \text{Art}\end{matrix}\right\} & \text{V} & \text{Case} & [+\text{pro}] & \longrightarrow & 1 \;\; \underline{\emptyset} & \overset{4}{[+\text{clit}]} & 2 \\ 1 & 2 & 3 & 4 & & & & \end{array}$$

This rule preposes actor and subject pronouns, and places them in the *0* case. It is therefore a kind of anti-passive, although it is restricted to pronouns (as in all Polynesian languages).

Finally, we should note that the analysis of Clitic Placement as an anti-passive is based on the fact that it has an accusative case structure. Like the clitic rule in Tongan, it affects pronominal actors and subjects, but not the objects of transitive verbs.

5.5 *Summary*. This brief survey of Clitic Placement rules suggests that they were once a general feature of Polynesian. Although Clitic Placement survives in only a handful of languages, it occurs in two of the three major subgroups and uses an old morphological set of pronouns. Several languages in the third, Eastern subgroup have a construction which seems to be descended from the Clitic rule. We can therefore reconstruct the rule for proto-Polynesian, as either applying to active (as in Maori) or passive (as in Pukapukan) sentences:[79]

(T9) Clitic Placement A (optional)

$$\begin{array}{ccccl} \text{Tns} & \text{V} & [\text{+pro}] & \text{NP} & \longrightarrow 1\ \ \underline{\text{a}}\ \ 3\ \ 2\ \ 4 \\ 1 & 2 & 3 & 4 & \end{array}$$

(T10) Clitic Placement B (optional)

$$\begin{array}{cccccl} \text{Tns} & \underbrace{\text{V (C)(i)a}} & \underline{\text{e}} & [\text{+pro}] & \text{X} & \longrightarrow 1\ \ \underline{\text{a}}\ \ 4\ \ 2\ \ 5 \\ 1 & 2 & 3 & 4 & 5 & \end{array}$$

Notice that these two versions have different ordering requirements, since CPA must precede the Passive, but CPB must follow it.

Despite the uncertainty in the reconstruction, it seems clear that Clitic Placement was originally an ergative rule. Moreover, it was affected by the accusative-to-ergative drift, since it is ergative in Pukapukan, but accusative in Samoan and Tongan. This shift in case seems to be designed so that the case structure of the rule is always the opposite of the sentences which it affects. We can therefore conclude that Clitic Placement obeys the revised restriction (C'):

> (C') An optional, clause-restricted rule may never have the same case structure as the basic case marking of the sentences to which it applies.

Because it obeys restriction (C'), Clitic Placement provides even more evidence that case is an organizing principle in Polynesian.

5.6 Anti-Passive or Not? Finally, the question remains of whether Clitic Placement can be considered an anti-passive in Tongan and Samoan. I have suggested that this equation might account for the fact that Possessive Preposing and Extraposition seem to be complementary (§4.5.5). However, the data of the preceding sections shows that the equation is not as simple as it seems.

An anti-passive shifts the focus of an ergative sentence by failing to stress the agency of the ergative noun. On a syntactic level, this shift is accomplished by placing the actor in the absolutive, and the object in a peripheral case. We have seen that Clitic Placement in Tongan and Samoan satisfies some of this definition. Clitic Placement shifts the emphasis away from pronominal agents by moving them to the second position in the sentence. Because this position is also allowed for pronominal subjects, but not objects, the rule could be said to place the agents in an absolutive position. In Samoan, some evidence suggests that this position is marked by the absolutive 0; for if it is not, the similarity between Clitic Placement and Pronoun Preposing would simply be an accident. The parallels between Clitic Placement and Possessive Preposing throughout Polynesian suggest that this cannot be the case.

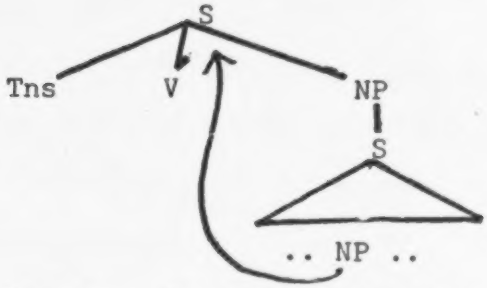
On the other hand, there are some aspects of Clitic Placement which do not satisfy the definition. An anti-passive should apply to all nouns which are actors, but Clitic Placement is limited to pronouns. An anti-passive should place the object of the sentence in a peripheral case,[80] but the object of a sentence like (324) remains in the absolutive. Finally, an anti-passive should resemble a passive rule in being a derived and special construction. This is not true for Clitic Placement in Tongan and Samoan, which is more likely to apply to ergative sentences than not.

These facts combine to suggest that Clitic Placement is not an anti-passive in the strict sense of the word; rather, it is a more restricted type of rule which manipulates word order and is limited to pronouns. Still, there is some truth in the observation that the rule functions like a passive in accusative languages. Clitic Placement patterns actors with subjects more clearly than any other clause-restricted rule in Tongan. Consequently, it seems fair to say that it is the closest that Polynesian comes to having an anti-passive.

6.0 A Two-Clause Rule (Raising)

The rules in §3-5 are all restricted to clauses, and thus fundamentally similar to case marking rules. I have suggested that this resemblance can account for the fact that both types of rules are radically affected by the accusative-to-ergative drift. If this suggestion is at all correct, we would expect rules which refer to more than one clause to behave rather differently. This prediction is borne out by the Raising rule, whose SD is apparently the same throughout Polynesian.

All four representative languages have a rule of Raising, which raises an NP to the subject position of a higher clause.[81] Raising usually operates out of sentential subjects, as illustrated on p. 129:



The details of Raising vary from language to language, but (with one possible exception) its structural description does not change throughout the accusative-to-ergative drift. What does change, however, are the types of NP's which are allowed to undergo the rule. The discussion which follows suggests that this variation is not an accident; instead, the NP's can be systematically defined by referring to the notions of case and logical subject.

6.1 Maori. Raising in Maori applies most often to negative sentences, which are formed by making the negated sentence into the subject of a higher negative verb. Compare the affirmative sentence (332) with its negative, (333):

(332) kua oma 0 te tamaiti. "The child ran."
tns run nom the child

(333) kaahore kia oma 0 te tamaiti.
neg tns run nom the child
"The child hasn't run yet."

Raising may optionally apply to (333), and turn the subject of the lower clause into the subject of the negative:

(334) kaahore 0 te tamaiti kia oma.
neg nom the child tns run

"The child hasn't run yet."

Notice that (334) cannot be explained as an instance of Equi Noun Phrase Deletion, since the subject in (333) clearly belongs to the lower clause.[82] Moreover, if (334) had resulted from Equi, we could not explain the form of the tense marker in (336). (335-6) show that whenever Equi applies, the tense marker kia must be changed to ki te:

(335) i hiahia 0 au kia kimi 0 a Maaui
tns want nom I tns look nom pers M.
i oona maatua.
acc his-pl parent
"I wanted Maaui to look for his parents."

But:

(336) i hiahia 0 au ki te kimi i ooku maatua.
*kia
tns want nom I tns look acc my-pl parent
"I wanted to look for my parents."

But negative sentences always take kia, and the ki te form is not allowed:[83]

(337) *kaahore 0 te tamaiti ki te oma.
not nom the child tns run

(337) shows that the NP following the negative must have originated in the lower clause.

Raising also applies to accusative sentences, where it affects the actor, but never the object NP:

(338) kiihai 0 a Pita i awhiti i a Mere.
neg nom pers P. tns help acc pers M.
"Pita didn't help Mere."

From:

(339) kiihai i awhiti Ø a Pita i a Mere.
neg tns help nom pers P. acc pers M.
"Pita didn't help Mere."

Raising operates on an actor NP in (338). (340) shows what happens when an object is raised instead:

(340) *kiihai Ø a Mere i awhiti (ai) Ø a Pita.
i
neg pers M. tns help rel nom pers P.

This pattern is reversed for passive sentences, where object NP's can undergo the rule:

(341) kaahore Ø a Mere i awhiti - a e Pita.
neg nom pers M. tns help pass agt P.
"Mere wasn't helped by Pita."

From:

(342) kaahore i awhiti - a Ø a Mere e Pita.
neg tns help pass nom pers M. agt P.
"Mere wasn't helped by Pita."

However, actor NP's cannot, as demonstrated by (343):

(343) *kaahore Ø (a) Pita i awhiti - a Ø a Mere.
e
neg pers P. tns help pass nom pers M.

These facts can be easily formalized when we remember that the NP's in (334, 338, 341) have two characteristics in common; all are preceded by the nominative Ø, and all were the NP's closest to their original verbs. (This second statement assumes, of course, that Raising is ordered before the Scrambling rule.) If we assume that the relevant characteristic is 'NP closest to the verb', then the rule can be stated as follows:

(TM9) Raising (optional)

$$\begin{array}{cccccc} V & [_{NP} & \underbrace{[_S\ \text{Tns}\ \ V} & \underbrace{\text{Case NP}} & \underbrace{X\ {}_S]} & {}_{NP}] \\ 1 & 2 & 3 & 4 & 5 & 6 \end{array}$$

$$\longrightarrow 1\ \ 4\ \ 3\ \ 5$$

(TM9) raises the NP next to the lower verb along with its case marker (which is always Ø). The case marker is included because raised NP's undergo later rules like the subjects of intransitives, or any nominative NP. For instance, Topicalization is optional for subjects of intransitives:

(344) ko te tamaiti kua oma. "It is the child
top the child tns run who ran."

And affects raised NP's in exactly the same way:[84]

(345) ko Pita kiihai i awhiti i a Mere.
top Pita neg tns help acc pers M.

"It is Pita who didn't help Mere."

Although (TM9) does not have a particular case structure, it is important that it applies to the NP's which are traditionally defined as grammatical subjects. These are the subjects of intransitives, actors of transitives, and objects of passive verbs, all of which are preceded by the nominative Ø.

6.2 Tongan. Raising in Tongan has the same description as Raising in Maori, but the paradigm which it generates is rather different. The rule still moves the subjects of intransitives, as in (346):

(346) 'oku lava Ø e tangatā 'o huu ki hono falē.
tns can abs the man compl go to her house

"The man can go into her house."

From:

(347) 'oku lava ke huu Ø e tangatā ki hono falē.
tns can tns go abs the man to her house

"It is possible for the man to enter her house."

It also affects the actors of transitive verbs:

(348) 'oku lava 'e Siale 'o hua'i Ø e hu'akaū.
tns can erg Siale compl throw abs the milk

"Siale can throw out the milk."

From:

(349) 'oku lava ke hua'i 'e Siale Ø e hu'akaū.
tns can tns throw erg Siale abs the milk

"It is possible for Siale to throw out the milk."

Both raised nouns in (346) and (348) keep their original case markers, and the other type of case is apparently not allowed. According to the informant,[85] a raised subject sounds "strange" in the ergative case:

(350) *?'oku lava 'e he tangatā 'o huu ki hono fale.
tns can erg the man compl go to her house

and a raised actor may not appear in the absolutive:

(351) *'oku lava Ø 'a Siale 'o hua'i Ø e
tns can abs pers S. compl throw abs the

hu'akaū.
milk

Sentences like these establish that the raised nouns do begin in the lower clause, and consequently that Raising is a genuine rule in Tongan.[86] This rule differs from Raising in Maori in two important ways. It is not limited to NP's in the Ø case, as shown by (348). And it never applies to objects of transitive or (lexicalized) passive verbs:

(352) *na'e lava Ø e fefinē 'o taa'i 'e Siale.
tns can abs the woman compl hit erg S.

Compare:

(353) na'e lava 'e Siale 'o taa'i Ø e fefinē.
tns can erg Siale compl hit abs the woman

"Charlie managed to hit
the woman."

(354-5) contains a lexicalized passive:

(354) *'oku lava Ø e fefinē 'o li'eki - na 'e
tns can abs the woman compl leave pass agt

he tangatā.
the man

(355) 'oku lava 'e he tangatā 'o li'eki - na
tns can erg the man compl leave pass

Ø e fefinē.
abs the woman

Notice that this is true even when the actor has undergone Clitic Placement:

(356) *'oku lava Ø e fefinē 'o ne li'eki - na.
tns can abs the woman compl he leave pass

(357) *na'e lava Ø e fefinē 'o ne taa'i.
tns can abs the woman compl he hit

Although these differences seem rather large, the Tongan rule can also be formalized as (TM8). This is because Scrambling cannot apply to most transitive sentences in Tongan, and thus the actors of transitives and subjects of intransitives will be the NP's closest to the verb:

(TT8) Raising (optional, = TM8)

$$\begin{array}{cccccc} V & [_{NP} & \underbrace{[_S\ \text{Tns}\ \ V} & \underbrace{\text{Case NP}} & \underbrace{X\ _S]} & _{NP}] \\ 1 & 2 & 3 & 4 & 5 & 6 \end{array}$$

$$\longrightarrow 1\ \ 4\ \ 3\ \ 5$$

Finally, we should note that the NP's which undergo the rule cannot be identified as grammatical subjects in Tongan. Rather, they must be logical subjects, since they include all grammatical NP's but the object of a transitive verb.

6.3 Pukapukan. So far we have seen that the input of Raising is affected by the accusative-to-ergative drift. The rule applies to grammatical subjects in an accusative language, but to logical subjects in an ergative one. This fact might lead us to suspect that in languages in the middle of the continuum, Raising would apply to some intermediate type of NP. This prediction turns out to be true for Pukapukan, as I will demonstrate here.

Raising in Pukapukan bears a strong resemblance to

Raising in Maori. The rule applies most often to negative sentences, which are formed with the negative verb kiai:[87]

(358) kiai na lele Ø te puuaapii.
neg tns run nom the teacher

"The teacher didn't run."

From:

(359) na lele Ø te puuaapii. "The teacher ran."
tns run nom the teacher

In (360), the subject of the negated sentence has been raised to the higher clause:

(360) kiai Ø te puuaapii na lele.
neg nom the teacher tns run

"The teacher didn't run."

Raising also is permitted for actors of accusative verbs, as demonstrated by (361):

(361) kiai Ø naa waawine na yii i naa atu.
neg nom the-pl wom. tns catch acc the bonito

"The women didn't catch the bonito."

From:

(362) kiai na yii Ø naa waawine i naa atu.
neg tns catch nom the-pl wom. acc the bonito

"The women didn't catch the bonito."

However, it is never allowed for accusative objects:

(363) *kiai Ø naa atu na yii (ai) Ø naa
i
neg the-pl bon. tns catch rel nom the-pl

waawine.
women

Since Raising can also affect the object of a passive verb, we might be tempted to conclude that it applies to grammatical subjects:

(364) kiai Ø te wenua na tiaki - na e Te Malo.
neg nom the land tns lead pass agt T.M.

"The country was not led by Te Malo."

From:

(365) kiai na tiaki - na Ø te wenua e Te Malo.
neg tns lead pass nom the land agt T.M.

"The country was not led by Te Malo."

But this is not entirely correct, since actors of passive verbs can also undergo the rule:

(366) kiai Ø i a Te Malo na tiaki - na Ø te
neg nom pers T.M. tns lead pass nom the

wenua.
land

"Te Malo didn't lead the country."

Notice that the raised actor in (366) is not preceded by its original case marker, but seems to be preceded by the nominative Ø. This Ø is necessary because raised NP's undergo later rules like any other nominative noun.[88] Moreover, the change in case does not suggest that (366) has undergone Equi instead of Raising. It is normal for movement rules to remove the e from agent nouns, as shown by (367):

(367) na tiaki - na Ø te wenua e Te Malo.
tns lead pass nom the land agt Te Malo

"The country was led by Te Malo."

When Subject Shift has applied:

(368) i a Te Malo na tiaki - na 0 te wenua.
pers T.M. tns lead pass nom the land

"Te Malo led the country."

Not:

(369) *e Te Malo na tiaki - na 0 te wenua.
agt Te Malo tns lead pass NOM the land

Sentences like (364-6) are particularly surprising, because they show that Raising can move the actor or object of a passive sentence. This choice is not allowed for accusative sentences, as illustrated by (363). While the reason for this difference is not entirely clear, it seems to stem from the fact that Raising must refer to some kind of subject. Both NP's in a passive sentence can satisfy different definitions of subject, since the actor is a logical subject and the object is preceded by 0 (= grammatical subject). In accusative sentences, on the other hand, the object is preceded by an overt i. As the object is neither a logical nor grammatical subject in this case, it is not allowed to undergo the Raising rule. Notice that the same is not true for actors of passive verbs, which need not be preceded by 0 to be considered (logical) subjects.

Raising therefore depends on the notions of logical and grammatical subject; it can only move NP's which are not direct objects, or are preceded by the nominative 0. I should note that this conclusion is not substantially affected by the way that ergative sentences undergo the rule.

As demonstrated by (370-2), Raising does not apply to ergative sentences at all:

(370) kiai na patu Ø te taane e Yina.
neg tns hit abs the man erg Yina

"Yina didn't hit the man."

But:

(371) *kiai Ø te taane na patu e Yina.
neg abs the man tns hit erg Yina

(372) *kiai $\frac{\underline{\emptyset}}{\underline{e}}$ (i a) Yina na patu Ø te taane.
neg pers Yina tns hit abs the man

This fact simply reflects the status of the ergative as a new and special construction in Pukapukan.[89]

If we assume that ergatives are excluded by a lexical redundancy rule, then the Pukapukan version of Raising can be formalized as follows. Either Raising is restricted to the NP closest to the verb, and passive sentences may optionally be scrambled before the rule applies. (Recall that passives must be scrambled anyway before Pronominalization, §1.6):

(TP9a) Raising I (optional)

$$\begin{array}{ccccccc} V & [_{NP} & \underbrace{[_{S}\ \text{Tns}\ \ V} & \text{Case} & NP & \underbrace{X\ _{S}]} & _{NP}] \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array}$$

$$\longrightarrow 1\ \ \underline{\emptyset}\ \ 5\ \ 3\ \ 6$$

Or, the rule is simply restricted to NP's preceded by the markers e or Ø:

(TP9b) Raising II (optional)

$$\begin{array}{ccccccc} V & [_{NP} & [_S \underbrace{Tns\ V\ X} & \left\{ \begin{matrix} e \\ \underline{\emptyset} \end{matrix} \right\} & NP & \underbrace{Y}\ {}_S] & {}_{NP}] \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array}$$

$$\longrightarrow 1\ \underline{\emptyset}\ 5\ 3\ 6$$

From the evidence given in this section, it is impossible to choose between these two transformations. However, both analyses agree with the fact that Raising refers to logical and grammatical subject in Pukapukan.

6.4 *Samoan*. In Samoan, Raising applies to sentential subjects which are introduced by the complementizer *'ona*. This process is complicated, and seems to involve two separate Raising rules: one for non-specific NP's (introduced by *se*) and another for specific NP's (introduced by *le*).[90] For convenience, I will consider only the simpler of these rules, which raises specific NP's to the subject position of a higher clause.

Raising is optional for the subjects of intransitives and the actor of any transitive verb. The raised NP is preceded by *e*, as shown by (373):

(373)	e	'ese	∅	le	vave	*e*	∅	tama	'ona
	tns	emp	nom	the	qu.	erg	pl	boy	compl.

a'oa'o	-(ina)	∅	le	lesona.[91]
learn	pass	nom	the	lesson

"The boys did learn the lesson quickly."

From:

(374) e 'ese Ø le vave 'ona a'oa'o -(ina)
tns emp nom the quick compl learn pass

e Ø tama Ø le lesona.
agt pl boy nom the lesson

"It was quickly that the
boys did learn the
lesson."

Raising has applied to the actor of a passive/ergative in (373). In (375), it applies to the actor of an accusative:

(375) 'ua leva e a'u 'ona va'ai 'i Saamoa.
*Ø
tns long erg I compl see acc Saamoa

"It's been a long time
since I've seen Saamoa."

From:

(376) 'ua leva 'ona va'ai Ø a'u 'i Saamoa.
*e
tns long compl see nom I acc Saamoa

(377) shows a subject of an intransitive which has undergone the rule:

(377) e tatau e Pela 'ona alu 'i Saavai'i.
*Ø
tns must erg Pela compl go to Savai'i

"Pela must go to Savai'i."

From:

(378) e tatau 'ona alu Ø Pela 'i Saavai'i.
*e
tns must compl go nom Pela to Savai'i.

"It is necessary for Pela
to go to Savai'i."

All of these NP's must be preceded by *e*, regardless of their function in the lower clause. This fact might suggest that the NP's actually originated in the higher clause, and that the complex sentences in (373, 375, 377) had undergone Equi instead of Raising. However, this analysis cannot be right, because all higher verbs which take *'ona* are subcategorized for one NP.[92] Thus *vave* can be modified by a personal subject (cf. 373):

(379) e vave *Ø* Ioane. "John is quick."
tns quick nom John

But it cannot occur in a transitive construction:

(380) *e vave *e* Mele *Ø* Ioane.
tns quick erg M. abs John

(381) *e vave *Ø* Mele *'i* a Ioane.
tns quick nom M. acc pers J.

(381) is only grammatical in the meaning, "Mary is quicker than John."

The NP's in (373, 375, 377) must therefore have originated in the lower clause, and undergone a Raising rule. This rule might account for the particle *e* in one of two possible ways. *E* might be inserted by the Raising rule, in place of the original case marker of the noun:

(TS8a) Raising (optional)

$$\begin{array}{ccccccc} V & [_{NP} & [_{S}\ \underbrace{Tns\ V\ X} & Case & NP & \underbrace{Y\ _{S}]} & _{NP}] \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array}$$

$$\longrightarrow 1\ \underline{e}\ 5\ 3\ 6$$

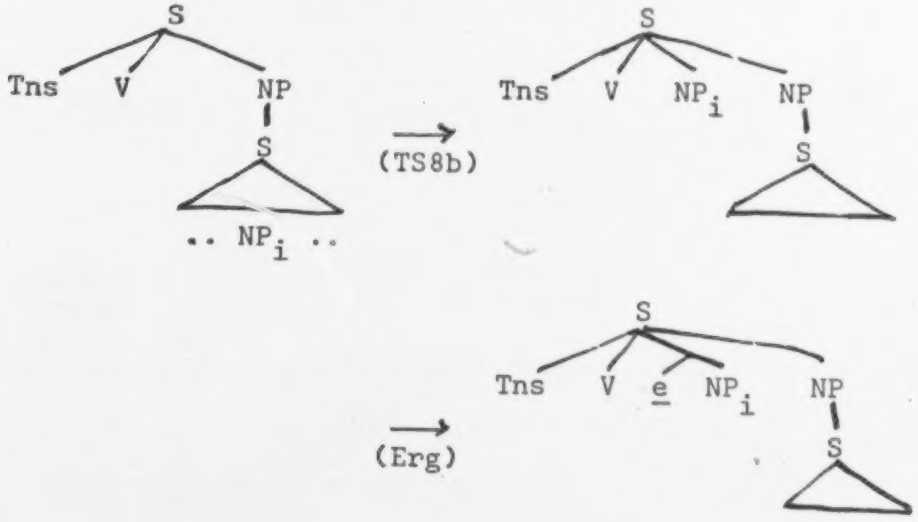
Or, Raising might eliminate the case marker of the noun entirely, but leave the complement as a constituent NP. If Ergative Marking were a cyclic rule, *e* would be inserted automatically in the resulting transitive construction:

(TS8b) Raising II (optional)

$$\begin{array}{ccccc} V\ [_{NP} & \underbrace{[_S\ \text{Tns}\ V\ X} & \text{Case} & NP & \underbrace{Y\ {}_S]}\ {}_{NP}] \\ 1 & 2 & 3 & 4 & 5 \end{array}$$

$$\longrightarrow 1\ \ 4\ \ 2\ \ 5$$

(TS8b) and Ergative Marking would operate as follows:



Although these analyses seem equally satisfactory, data from se-Raising suggests that the first solution is probably correct (v. fn. 90). We may then tentatively assume that Raising can be formalized as (TS8a).

Raising does not affect the object of a transitive verb, no matter what case marker precedes it in the higher clause. (382) shows the rule applying to the object of an accusative verb:

(382) *e mafai 'i Ø tamaiti 'ona angaleanga
e
Ø
tns can pl child compl ill-treat

(iai) 0 le tamaaloa.
rel nom the man

Compare:

(383) e mafai e le tamaaloa 'ona angaleanga
tns can erg the man compl ill-treat

'i 0 tamaiti.
acc pl child

"The man can mistreat
the children."

From:

(384) e mafai 'ona angaleanga 0 le tamaaloa
tns can compl ill-treat nom the man

'i 0 tamaiti.
acc pl child

"It is possible for the man
to mistreat children."

In (385), Raising affects the object of a passive/ergative:

(385) *e mafai 0 le talo 'ona totoo e laatou.
e
tns can the taro compl plant erg them

Compare:

(386) e mafai e laatou 'ona totoo 0 le talo.
TNS can erg they compl plant abs the taro

"They can plant the taro."

The fact that Raising applies to subjects but not objects was emphasized by my informants, whose first reaction to (385) was uncontrollable laughter. They claimed that (385) tried to describe "a plant doing something to a person," and sounded like "a different language."

As in Tongan, the fact that Raising does not operate on objects can be formalized by restricting it to the NP closest to the verb. (This assumes that Scrambling is ordered after the Raising rule.) The rule then looks exactly like Raising in Tongan

(TS8a) Raising I, revised (optional)

$$\begin{array}{ccccccc} \text{V} & [_{\text{NP}} & \underbrace{[_{\text{S}} \text{ Tns } \text{ V}} & \text{Case} & \text{NP} & \underbrace{\text{X}\ {}_{\text{S}}]} & {}_{\text{NP}}] \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array}$$

$$\longrightarrow 1\ \ \underline{e}\ \ 5\ \ 3\ \ 6$$

except that its SC also inserts the ergative particle e.

Finally, we should note that at least this type of Raising is restricted to NP's which are logical subjects. Raising in Samoan can affect any subject or actor, as in Tongan, but can never apply to the object of a transitive verb.

6.5 Summary. We have seen that the structure of Raising in Polynesian is not substantially affected by the accusative-to-ergative drift. In Maori and Tongan, the rule raises an NP along with its case marker in the lower clause. In Pukapukan and Samoan, this case marker is replaced by either Ø (in Pukapukan) or e (in Samoan). Nonetheless, the rules in all four languages are restricted to the NP closest to the verb. Because this structural description does not change, Raising differs fundamentally from the clause-restricted rules in §3-5.

What does seem to be affected by the drift are the types of NP's which can undergo the rule. We have seen that

Raising in Maori applies to any grammatical (or nominative) NP. In Tongan and Samoan, on the other hand, the rule is limited to logical subjects. The fact that this difference is conditioned by the drift is suggested by Pukapukan, where grammatical and logical subjects can undergo the rule. The only NP which is excluded by this definition is the object of an accusative sentence, which is preceded by *i*.

The reason behind this shift can be explained if we remember that case marking is also connected to the notion of logical subject. Accusative case marking reflects the notion of logical subject directly, as it marks the subject of an intransitive like the actor of a transitive, active verb. In ergative case marking, on the other hand, the subject is identified with an object, and the relation of logical subject is not expressed. This difference is apparently responsible for the fact that Raising is restricted to logical subjects in ergative languages:

> (D) In languages where logical subject is not reflected by the normal case marking, it must be reflected by the output of a rule like Raising.

(D) obviously does not refer to clause-restricted rules, since it is violated by Extraposition in Tongan, and Possessive Marking (*o*) in Samoan. The constraint can be made slightly more explicit by comparing Raising with other two-clause rules. For instance, Raising and Equi both refer to logical subject in Tongan; Relative Clause Formation and Topicalization do not. Since topicalized nouns have no case, and the case of a head noun does not interfere with Relative Clause Formation, the motivating principle seems to be:

> (D') If logical subject is not reflected by the (normal) case marking of a language, it must be reflected by a two-NP rule (where a two-NP[93] rule classifies NP's in more than one clause of a complex sentence).

(D') clearly applies to Raising, which always makes the raised NP into a subject, and no other kind of grammatical NP. It also applies to Equi, which only operates from higher subjects (and possibly, objects) in Polynesian.

While the reason behind this constraint is not entirely clear, it does demonstrate that logical subject is an important relation in Polynesian. (D') is clearly designed to keep the logical subjects from being confused, when they are not indicated already by the case marking of simple sentences. However, the constraint does not imply that accusative case marking is more basic than the ergative type. This is because Raising in accusative languages does not refer to logical subject, and thus may identify the object of a passive with the subject of an intransitive verb.

Consequently, logical subject cannot be identified entirely with one type of case marking. Rather, it has the complementary relationship with case which is expressed by (D'). This relation can be diagrammed as follows,

	Logical Subject in:	
Reflected by:	Accus. Lg.	Erg. Lg.
Case Marking	yes	no
Two-NP Rules	not nec.[93]	yes

and seems to be observed in many languages of the world. Thus Equi in an ergative language like Eskimo is restricted to logical subjects; Raising in English, an accusative language, does not have this restriction. Evidence of this sort does suggest that it is entirely possible that (D') (or some version of it) is a universal. If so, it would simply confirm that the logical subject relation is more basic than either type of case.

Footnotes

[*] Research on this paper was supported in part by the U.S. Department of Health, Education and Welfare, Office of Education, Grant OEC-0-70-4986(823) to the Language Research Foundation. Fieldwork in Auckland, New Zealand was sponsored by the Alice F. Palmer Travelling Fellowship (Wellesley College) and the Isobel Briggs Travelling Fellowship (Radcliffe College). A rough version of §2 was read at the first meeting of the Linguistic Society of the University of Auckland, 1971. I would like to thank my informants: Dr. Patrick W. Hohepa (Maori); Tupou Tonutonu, Tami Matheson (Tongan); Atawua Robati, Parepano T. Tukia (Pukapukan); Lucy Feaver, 'Ati, John Faumuinā (Samoan). Special thanks to Stephen Anderson, Alan Timberlake and Robert Underhill for criticisms and encouragement.

1. Suggestions of this sort have been made by Hale (1968) for Australian languages; by Hohepa (1969), Lynch (1971) and Anderson (ms.) for Polynesian.

2. This hypothesis was first suggested by Hale (1968) and Hohepa (1969).

3. The few exceptions belong to a closed class of stative verbs in Maori, which have been discussed by Hohepa (1970). Statives are active verbs with semantically passive meanings, e.g. mate 'die', mamae 'be wounded'. The objects of these verbs appear in the nominative, while the actors are preceded by an 'agentive' i:

(a) i mate Ø te tangata i te wahine.
tns die nom the man agt the woman

"The man died because of the woman."

It is interesting that this i is homophonous with the accusative i, and not with the e that marks the agent of a passive.

Statives cannot undergo the passive rule, as illustrated by (b):

(b) *i mate - hia Ø te tangata e te wahine.
tns die pass nom the man agt the woman

This seems to agree with the fact that their normal case marking is ergative instead of accusative; that is, the object of a stative verb is identified with the subject of an intransitive. Why Maori should have this restricted type of ergative case marking is unclear to me. But it is significant that the class of statives is progressively reduced in the course of the accusative-to-ergative drift. In Pukapukan, adjectives which are stative in Maori have been reanalyzed as passives; compare makariri 'cold' in Maori with Pukapukan makalili-a 'cold'. In Tongan and Samoan, the class of statives is reduced to a mere handful of verbs. We can guess that this reduction is motivated by the appearance of a productive ergative construction in these languages.

4. Note that moe is intransitive, not transitive, since the locative prepositional phrase may be omitted:

(a) ka moe Ø te tangata. "The man slept."
tns sleep nom the man

(A) means "the man slept", and not "the man slept on something."

5. See Hale's review (1968a) of Hohepa.

6. Shumway (1971), p. 140. Tongan is practically the only Polynesian language where *'a* occurs as an absolutive marker. The evidence of all the other languages points to the fact that *a* was originally a marker for pronouns or personal names. This marker appeared in the nominative (*Ø*) and accusative (*i*) cases. In Tongan, *'a* was apparently reanalyzed as a marker of the absolutive (formerly nominative), and then generalized to all nouns. This development might be represented:

$$\underset{\text{abs}}{\underline{\emptyset}}\ \underset{\text{pers}}{\underline{\text{'a}}}\ \text{PropN} \longrightarrow \underset{\text{abs}}{\underline{\text{'a}}}\ \underset{\text{pers}}{\underline{\emptyset}}\ \text{PropN} \longrightarrow \underset{\text{abs}}{\underline{\text{'a}}}\ \text{(Art) N}$$

It is interesting that the most recent developments in Tongan seem to reverse this change, since Shumway states that *'a* is eliminated before common nouns (but not pronouns or proper names). There are two possible explanations for this fact. Either the change described by Shumway eliminated *'a* when it was the first of two particles preceding the noun; in this case, the change is a simple reversal. Or, it might be that on a deeper-than-surface level, the *'a* before common nouns and that before proper nouns are not the same. This explanation avoids reversing the first change, but is suspicious on a descriptive level.

7. *'i* is eliminated everywhere but before pronouns and proper names, which also take the personal *'a*: *'i* *a* *Mele* 'to Mele'. The fact that both particles survive here is clearly not an accident.

8. But see below, (52-3).

9. The fact that most passive verbs cannot be accusative suggests that passive sentences should not be derived by means of a passive rule. Instead, they should be derived by a low-level suffix attachment rule from ergative sentences (which they resemble in word order, distribution and other properties). Consider:

(Ta) Ergative Marking

$$\begin{array}{lll} V & NP & NP \\ 1 & 2 & 3 \end{array} \longrightarrow 1\ \underline{e} + 2\ \underline{Ø} + 3$$

(Tb) -$\underline{(C)(i)a}$ Attachment (optional)

$$\begin{array}{ll} V & \underbrace{\underline{e}\ NP} \\ 1 & 2 \end{array} \longrightarrow 1 + \underline{(C)(i)a}\ \ 2$$

Passive verbs which can also be accusative might be derived by an ordinary passive rule; or, they might be listed as separate lexical items (with an obligatory Tb). This second solution is compatible with the fact that passive forms of accusative verbs are often assigned slightly different meanings. Compare *vala'au* 'call (acc.)' and *vala'au'lia* 'invite (pass.)'.

10. This generalization is not strictly true, since several accusative verbs have been reanalyzed as ergative ones in present-day Samoan. Often this reanalysis is accompanied by a change of meaning; e.g. *va'ai* 'inspect, see (a patient) (acc.)', but *va'ai* 'see (erg.)'.

11. Pronominalization works from left to right in all Polynesian languages of which I am aware. My description of pronominalization states that the anaphoric NP's are replaced by a pronoun, but I take no stand on the theoretical issue of how pronominalization works (e.g. by NP replacement, referential indices, etc.), and the choice of one theory of pronominalization over another does not seem to affect the analysis.

12. Chapin (1970) makes the claim that Samoan allows left-to-right and right-to-left pronominalization in coordinate structures. This is simply denied by my three Samoan informants, who insisted that (62) was not co-referential. It seems likely that Chapin's Samoan informant did not know his language that well, since he also approved Samoan sentences with split conjuncts (v. Grinder, 1970). Other Samoans and scholars of Samoan (for instance, Andrew Pawley) are of the opinion that split conjuncts and right-to-left pronominalization are simply not part of the language.

13. Strictly speaking the Passive Rule should not be optional, since it changes meaning (e.g. the verb is often completed or resultative) and focus (from actor to object). But I take the view that these changes of meaning could be accounted for by later, interpretive rules; so a constituent by-manner, or whatever, would be unnecessary in the deep structure.

14. I will not consider pronominalization in Samoan

(which works exactly like pronominalization in Pukapukan) because the passive rule is not productive in Samoan. Instead, passive sentences are derived for the most part from ergative sentences by a -*(C)(i)a* Attachment rule.

15. Thus the pronouns in (79-80) are derived by primacy (actor-to-object); those in (81, 79) are derived by precedence (left-to-right). Notice that the two types of pronominalization can both be said to work on a sentence in the normal VSO order (79). I am unclear as to how the constraint on scrambling should be formalized, but something like it is clearly necessary to explain the Maori data.

16. It is interesting that ergative sentences in Pukapukan have exactly the same pronominalization pattern as passive ones. Consequently, pronominalization does *not* provide evidence against deriving the ergative from the passive. Consider:

(a) na patu *Ø* te *taane* e *na* wua. "The *man* hit
tns hit abs the man erg him emp *himself*."

(b) na patu e te *taane* Ø i a *ana* wua.
tns hit erg the man abs pers him emp

"**Himself* hit the *man*."

(A) and (b) are coreferential; (c) and (d) are not.

(c) na patu e na wua Ø te taane. "Someone hit
tns hit erg him emp abs the man the man."

(d) na patu Ø i a ana wua e te taane.
tns hit abs pers him emp erg the man

"The man hit someone."

17. On the other hand, if (83-4) are derived by precedence,

we are left with the problem of how they should be derived. Either Scrambling could apply to sentences before Pronominalization, and some accusative sentences would be prevented from undergoing the rule. (Recall that in Pukapukan, accusative sentences containing pronouns or proper names cannot be scrambled at all.) Or, the Passive might reorder actor and object only optionally, so that sentences with both word orders would be produced. Both of these solutions have their disadvantages, but I believe that the first is probably closer to the truth. In this case, Scrambling will have to be a "peeking rule", since it will have to know which accusative sentences will be pronominalized (and thus should not undergo Scrambling).

18. That is, Scrambling will be obligatory for most sentences in Tongan, but optional in Maori.

19. This is the analysis of Hohepa (1969). Hohepa considers and rejects an alternative analysis, in which the passive suffix is -0 in ergative languages. I do not understand completely his reasons for rejecting a passive -0, but I feel that either analysis would not affect the substance of the H/H proposal.

20. Hohepa (1969), p. 325.

21. Hale (1968) claims that one of the main advantages of his analysis is that it reduces the accusative-to-ergative drift to a change in conditions on a lexical redundancy rule. Accusative languages will contain a redundancy rule stating

that the passive is optional for transitive sentences. Ergative languages will have lost this redundancy rule, so that the passive will be obligatory for transitive sentences. According to Hale, this will be the only major difference between related accusative and ergative languages.

22. These details could be accounted for by some lexical redundancy rules:

[+emot/percep] ⟶ [±Passive]

[+emot/percep] ⟶ [- (C)(i)a Delet]

23. The particles preceding the nouns in Niuean are combinations of case particles and the definite article. Historically, they seem to be descended from a case system like that of Tongan, with the case particles *'e* and *'i* being eliminated to parallel the absolutive *0*. (Note that this development is possible because the definite articles are different for absolutive and ergative or accusative nouns. It is also possible because word order in Niuean is strictly VSO.)

24. This is especially obvious in languages where the passive can apply to intransitives followed by a locative phrase. Thus in Pukapukan,

(a) na lele 0 te taane. "The man ran."
tns run nom the man

means "the man ran", with no implied direct object. In (b), however,

(b) na lele - wia 0 te taane. "The man was run over."
tns run pass nom the man

the verb is transitive, because both an actor and an object are implied. ("The man was run over (by something)" is the complete translation.)

25. Technically speaking, passive sentences should be intransitive, since the actor has been removed to a peripheral case. It seems that this cannot be true in ergative languages, however. If it were, ergative languages would have no transitive sentences at all. This observation has led Hohepa to assume that passive sentences are reanalyzed as transitives in ergative languages.

26. ke is a clitic pronoun which is not preceded by any overt case particle. The fact that the sentence is accusative is demonstrated by the 'i preceding the object (which can be omitted in casual speech).

27. What seems to happen in the course of this lexicalization is that the passive suffix takes on several different meanings, all of which can be traced back to the syntactic passive. Thus in Tongan the passive suffix can indicate completed action, result, duration, or intensity; these are nuances of the passive in accusative languages. The Tongan passive can also indicate that the logical subject of the verb is not the semantic agent of the action; in accusative languages, this meaning is obtained through an optional rule of Agent Deletion. Finally, the Tongan passive can indicate that its absolutive noun is different from that attached to the active verb. In accusative languages, this is accom-

plished directly by the mechanism of the passive rule.

28. Negatives are higher predicates in Polynesian languages; see Chung (1970) and Clark (1971) for discussion.

29. *'o* is a complementizer in Tongan which appears after certain higher verbs when Raising (and perhaps, Equi) has applied.

30. These sentences were elicited from Tevita 'Ahokovi by Ross Clark.

31. Since Clitic Placement produces sentences with a distinct morphological set of pronouns, it is probable that the rule should not be optional, but that pronouns should be specified for the feature [clitic] and those with the positive value of the feature would undergo the rule. I have chosen the representation on p. 43 because it is easier.

32. Strictly speaking, the object of a passive sentence should be placed in the nominative by the accusative marking rule. But the passive sentence is still "marked for case" by the passive rule, because the actor is preceded by *e* (and the object may not take the *i* of accusative marking).

I should mention in passing that my picture of the Passive as inserting a nominative *Ø* is merely a simplifying device; it is easier to think of passive sentences, and talk about them, as being derived in a totally different way than active accusative ones. This simplifying device does not affect any of the arguments in any way (for instance, the Double Case Marking argument could be restated as the claim that no sentence should have the passive *e* and the accusative i at once.)

33. An alternative solution might be to derive double case marking as a variant of the Ergative Marking rule. This rule would insert an optional i before objects of transitive complements. Then, sentences of the form V e NP i NP would be reordered to V i NP e NP. This analysis has the advantage of relating ergative case marking to double case marking; it has the disadvantage of requiring an obligatory Scrambling rule.

34. Cf. fn. 13.

35. The two forms of the verb do have different meanings; va'ai (acc.) means "to examine (as a doctor examines his patient;" va'ai (erg.) simply means "to look at."

36. For instance, Anderson (1968).

37. If we assume that much of historical change has to do with rule introduction, reordering, simplification, generalization and reanalysis.

38. Exceptions to this generalization are Niuean and Luangiua, both of which have recently eliminated one of the two types of possessive.

39. Scrambling may not affect nominalizations at all, as demonstrated by the following examples. (A-b) show that Scrambling cannot apply to nominalizations before possessive marking:

(a) *kua whakawaa - tia Ø te patu - nga o te
a
tns investig. pass nom the hit nomin poss the

wahine Ø Hoone.
woman nom Hoone

(b) *kua whakawaa - tia Ø te patu - a - nga a Hoone
o
tns investig. pass nom the hit pass nomin poss J.

Ø te wahine.
nom the woman

The other examples show that Scrambling cannot apply to nominalizations in any other case:

(c) *kua whakawaa - tia Ø te patu - nga i te
tns investig pass nom the hit nomin acc the

wahine a Hoone.
woman poss John

(only in the meaning:
"The killing of John's
woman (by someone)...)

(d) *kua whakawaa - tia Ø te patu -(a)- nga
tns investig. pass nom the hit pass nomin

e Hoone o te wahine.
agt John poss the woman

(e) *kua whakawaa - tia Ø te patu - nga e Hoone
tns investig. pass nom the hit nomin agt John

i te wahine.
acc the woman

(A-b) should be compared with (150) and (157); (c-e) should be compared with (150), (156), (157), (155).

40. Other languages have similar constraints which state that a complex sentence should have one and only one nominative NP; e.g. Latin (where subjects of infinitives are placed in the accusative), Southern Paiute.

41. Consider, for example

(a) *kua whakawaa - tia Ø te patu - nga a Hoone
tns investig. pass nom the hit nomin poss John

o te wahine.
poss the woman

Other examples with double possessives give similar results.

42. Evidence from Samoan, Pukapukan, Tongan, Niuean suggests that Preposing was originally limited to pronouns. None of these languages allow preposing for full noun possessors: Samoan and Pukapukan allow it optionally for pronouns. (In Tongan and Niuean, Preposing is obligatory for pronouns.)

43. See Churchward (1953), §30.7.

44. See Churchward (1953), §30.5.

45. See Churchward (1953), §30.11.

46. Thus sentences like the following are ungrammatical:

(a) *'oku 'ikai tonu Ø 'e - ne taki 'e he tu'ī.
tns neg right abs poss him lead erg the king

Compare (191).

47. The different forms of the possessive in (171) --'e and ho-- result from the fact that definite articles and possessive pronouns have been merged in Tongan.

48. Apparently, Subject Possessive Marking is not allowed at all for subjects of stative verbs; or so it would seem from an inspection of Shumway (1971). Shumway's examples of nominalizations with the stative mate 'to die' never show the subject of the verb in the possessive. On the other hand, non-statives like 'alu seem to allow their subjects to be possessivized optionally (v. Churchward, 1953).

49. This argument does not apply to pronominal actors,

which must undergo Possessive Preposing, and thus are never preceded by e in surface structure.

50. R. Underhill has suggested to me that the reason for the restriction might be that Possessive Marking was never meant to apply to agents of passive sentences. Presumably, the e case is an adverbial one, and the rule is limited to grammatical cases like the nominative or accusative. This restriction survives even in ergative languages, where e has been reanalyzed as a grammatical case (the ergative), and the restriction has no synchronic motivation.

51. Notice that (185) may not be scrambled again, as is shown by (187). The Extraposition rule is clearly a type of scrambling which is limited to nominalizations.

52. Notice that Possessive Marking for Objects does not affect the subjects of intransitives (which are already in the possessive) or the actors of transitives (which are preceded by 'e).

53. Chapin has claimed (1970) that Samoan has emphatic copies, but this is denied by my informants for the 3rd pers.

54. Notice that Scrambling does not apply to nominalizations at all. It does not apply before Possessive Marking o or (a):

(a) *e mea kino Ø te patu - nga o te wawine
a
a thing bad nom the hit nomin poss the woman

Ø i a Turi.
nom pers Turi

(b) *e mea kino Ø te patu - (a) - nga a/o Turi
a thing bad nom the hit pass nomin poss T.

Ø te wawine.
nom the woman

Nor does it apply in any other case:

(c) *e mea kino Ø te patu - nga i te wawine
a thing bad nom the hit nomin acc the woman

a Turi.
poss Turi

(except in the meaning:
"The killing [of] Turi's
woman (by someone)...")

(d) *e mea kino Ø te patu -(a)- nga e Turi
a thing bad nom the hit pass nomin agt Turi

o te wawine.
poss the woman

These examples should be compared with (197), (199), (212). The following examples should be compared with (204) and (211):

(e) *e mea kino Ø te patu - nga Ø te wawine
a thing bad nom the hit nomin abs the woman

e Turi.
erg Turi

(f) *e mea kino Ø te patu - nga e Turi
a thing bad nom the hit nomin agt Turi

i te wawine.
acc the woman

55. Notice that ergative nominalizations always occur in the order VSO (v. 204), and sentences like the following are not allowed:

(a) *e mea kino Ø te patu - nga Ø te wawine
a thing bad nom the hit nomin abs the woman

e Turi.
erg Turi

On the other hand, passive nominalizations are always in the order VOS:

(b) *e mea kino Ø te patu - a - nga e Turi
a thing bad nom the hit pass nomin agt Turi

o te wawine.
Ø
the woman

These differences in the form of passive and ergative nominalizations provide more evidence that the two constructions come from different sources in Pukapukan.

56. This rule will state that NP's preceded by a will be interpreted as dominant, active, and subject-like; NP's preceded by o will be interpreted as neutral, passive, and object-like. The rule can also be generalized to ordinary possessive constructions, where the two types of possessive have essentially the same meanings.

57. Example (f) in fn. 54 shows that Extraposition must reorder the actor and object.

58. This preference was voiced by my informants, and is supported by Churchward (1951), §418.2.

59. Notice that the same is not true of an actor which is in the possessive, since sentences like the following are ungrammatical:

(a) *e lelei Ø le aumai -(ina) Ø le tusi a
o
tns good nom the bring pass abs the book poss

Ioane.
John

(except in the meaning:

"the delivering (by some-
one) of John's book...")

(b) *'ua leva Ø le lee va'ai 'i a maatou
tns long nom the not see acc pers us

a le fooma'i.
o
poss the doctor

60. Example (225) shows that Possessive Marking is optional for non-singular clitics. The rule is required for clitics which are singular, as demonstrated by (a):

(a) *pe 'e te fiafia 'ona 'o le Ø na alofa
ia
Q you tns happy becaus. pred the abs he love

'i le teine.
acc the girl

It is interesting that ergators which have not undergone Preposing may only be placed in the a form of the possessive. Preposed clitics, on the other hand, may take either a or o. Compare (242-3) with the following:

(b) Ø 'aua lua te popole 'i le ta'e - ina
tns neg you tns worry acc the break pass

a naia o le fangu.
*o
poss him poss the glass

"Don't you-two worry about
the breaking of him of the
glass (i.e. his breaking
of..)"

(b) shows that it is false to say that clitics are possessivized like subjects of intransitives because they are pronouns, and not ordinary full nouns. The correct version of the relationship is that clitics act like subjects because they have undergone the Preposing rule.

61. Sentences (245-6) are derived from:

(a) 'ua ulaula-tapa'a Ø le faife'au.
tns smoke tobacco nom the minister

"The minister does tobacco-smoking."

(247-8) are derived from:

(b) 'ua faitau-tusi Ø 'oia. "He does book-reading."
tns read book nom he

(249-50) are derived from:

(c) 'olo'o mana'o Ø a'u iai. "I want it."
tns want nom I rel

62. Preposing must therefore precede either of the Possessive Marking rules.

63. Cf. (225), where the clitic need not be placed in the possessive at all.

64. The parenthesis notation in (TS3) is meant to indicate disjunctive ordering; the SD applies as e NP Y Ø NP or Ø NP Y Ø NP, before it applies as Ø NP Y Z or e NP Y Z.

65. (255) is only grammatical if it means "the bringing of Pita's book (by someone)..".

66. Possessive Marking must refer to the leftmost NP because it must refer to NP's before and after the verb, at the same time. This is because Pronoun Preposing must precede the possessive rules, and Possessive Marking must treat clitics and full-noun subjects in exactly the same way.

67. Lithuanian, for instance, has a single possessive form for all NP's in a nominalization. V. Senn (1966).

68. I consider the Samoan passive to be derived from an ergative construction.

69. See the chart on p. 104 for the ordering of Extraposition. The Maori passive is ordered before Accusative Marking, Pronominalization, Topicalization, Imperative Formation, *he*-Insertion, Subject-Raising, among others.

70. It is interesting that the loss of the nominalizing suffix in Nominalization Formation can also be attributed to the accusative-to-ergative drift. As we have seen, the development of -*(C)(i)a* and -*(C)(a)nga* is similar in Polynesian, and it is possible that the latter was eliminated in ergative languages on the analogy of the former.

71. A look at the available material on Polynesian suggests that Uvean, Futunan and Tokelauan are the only other languages where Clitic Placement survives. I am indebted to Bruce Biggs for the information on Futunan, and Peter Sharples for the information on Tokelauan.

72. An interesting reanalysis of the pronouns has taken place in Pukapukan, where singular full pronouns are used in the nominative (*0* case), and singular clitics are used in preposed position or after prepositions. Cp. *0* *au* 'I (nom.)' and *e* *ku* 'I (erg.)'. Dual and plural clitics have been replaced everywhere by the full pronouns.

73. The fact that the plural pronouns originally had clitic forms is supported by evidence from Tongan (Churchward, 1953).

74. Languages which meet this description include Hawaiian, Tahitian, Rarotongan and Maori, among others. Pukapukan has both Clitic Placement and the *naa*-construction, but does not provide a counterexample because it has borrowed the *naa*-construction from Rarotongan.

75. Once the reanalysis has occurred, of course, the *naa*-construction is separate from Clitic Placement, and may change independently of it. Thus in Maori, a future *maa*-construction was developed on analogy with *naa*-:

(a) maa te tangata e patu Ø te poaka.
poss the man tns hit nom the pig

"*The man* will hit the pig."

The *maa*-construction is not found in Rarotongan, Tahitian or Hawaiian, and thus seems to be an innovation of Maori. However, it is important that the *m*-/*n*- contrast suggests that *n*- originally marked a non-future tense.

Rarotongan has developed a *noo*-construction which is not found in other languages (e.g. Hawaiian). This construction preposes nouns which indicate non-active responsibility, and places them in the *o* possessive. These nouns are not grammatical NP's and must therefore be copied by the relative *ei*:

(b) noo Tumu i 'anau mai ei Ø a Tii.
poss Tumu tns born here rel nom pers T.

"Because of Tumu, Tii was born."

Compare:

(c) i 'anau mai Ø a Tii. "Tii was born."
tns born here nom pers T.

This extension follows naturally from the fact that the *a* and *o* possessives are opposed in Rarotongan.

Finally, the construction has been extended in Maori to indicate indirect agency of any sort. This indirect *naa*-construction may apply to sentences of any type, and affects the indirect agent (normally preceded by the agentive *i*):

(d) naa Hoone i mate ai 0 te wahine.
poss Hoone tns die rel nom the woman

"Because of John, the woman died."

(D) is derived from the stative sentence:

(e) i mate 0 te wahine i a Hoone.
tns die nom the woman agt pers John

"The woman died because of John."

Notice that the agent in (d) must be copied by the relative marker *ai*:

(f) *naa Hoone i mate 0 te wahine.
poss Hoone tns die nom the woman

Sentences like (d) seem to be the Maori equivalent of the *noo*-construction in Rarotongan.

76. I should note that for Possessive Preposing, this favoring of subjects seems to involve a surface structure constraint. Pronominal subjects are possessivized (and preposed) in preference to pronominal objects if both constituents will appear in surface structure (189-90). If, however, the subject will be deleted, the object must be possessivized instead:

(a) 'oku 'ikai tonu Ø ho - no taki o'o - na.
tns not good abs poss his lead poss his

"His being led is not good."

Not:

(b) *'oku 'ikai tonu Ø e taki Ø ia.
tns not good nom the lead abs him

Since the deleted subject in (a) must be a pronoun in underlying structure, the sentence cannot be accounted for by the rules in §4.2. (Notice that the subject cannot be deleted before Possessive Marking applies, because then the object would be placed in the a form of the possessive.) Instead, Subject Possessive Marking must be modified in this way:

> If the subject pronoun will be phonologically Ø in surface structure, Subject Possessive Marking may not apply and Extraposition is required.

77. Even as revised, restriction (C') seems a little simplistic, so it is interesting that there is at least one possible counterexample to it in Polynesian. Grezel (1878) states that Clitic Placement in Futunan affects the actor of a transitive and the subject of an intransitive verb:

(a) na kau nofo iai. "I stayed there."
tns I stay rel

(b) na ke maua i fea Ø lau sele.
tns you find where abs your knife

"Where did you find your knife?"

The rule seems to be optional, since Grezel also has:

(c) na fano Ø a ia. "He went."
tns go nom pers he

(d) kuo soli e au Ø loku toki.
tns give erg I abs my axe

"I gave him my axe."

(C-d) show that the basic case marking in Futunan is ergative.

However, more recent evidence from Biggs (1971) seems to indicate that Clitic Placement has been reanalyzed to apply to the absolutive case. Consider Biggs' examples:

(e) na kau ano. "I went."
tns I go

(f) na kau tiaki e loku tinana.
tns I reject erg my mother

"I was rejected by my mother."

Since Biggs gives no examples of pronouns aside from these, it is impossible to tell whether or not the rule is still optional.

(E-f) present an interesting problem, for they suggest that in the space of 100 years Futunan has completely changed the case structure of its clitic rule. But if this change actually occurred, and if Clitic Placement is still optional, it would constitute a counterexample to (C').

I should mention that rules like Topicalizations are not obvious counterexamples to (C'), because their SC moves the topicalized NP out of its original clause. For this reason, I do not consider them to be clause restricted rules.

78. Alternatively, it might be possible to have clitic placement insert 0 before the preposed pronoun. Then, a rule would insert the personal a before objects of accusative sentences, and any nominative pronoun or proper name. This version of the rule would work slightly better if the personal marker had not been reanalyzed as i a before nominative full nouns.

79. Notice that CPA must precede Accusative Marking.

80. The anti-passive in Eskimo, for instance, places the actor of a transitive in the absolutive case, and the object in the instrumental.

81. This section is only concerned with Subject Raising to subject position, although some Polynesian languages (e.g. Tongan) may allow Raising to object position as well.

82. This is because complement sentences are usually extraposed to the right in Maori.

83. Moreover, the raised NP in (334) must be the subject of the negative, and not a topicalized form. This fact is established by the following evidence.

Maori has a non-specific article he, which may precede nouns in certain positions in the sentence. It may modify the subject of an intransitive, as in (a). But it may not occur with the actor or object of a transitive (accusative) verb:

(a) i oma 0 he tangata. "A man ran."
tns run nom a man

(b) *i patu 0 he tangata i te poaka.
tns hit nom a man acc the pig

(c) *i patu 0 te tangata i he poaka.
tns hit nom the man acc a pig

Since *he* is ungrammatical in (c), it is interesting that it can modify the object (= superficial subject) of a passive verb.

(d) i patu - a 0 he poaka e te tangata.
tns hit pass nom a pig agt the man

"A pig was killed by the man."

(D) suggests that passives are a type of intransitive in Maori; it also suggests that *he* is inserted after the Passive rule has applied. Otherwise, if *he* were introduced in the deep structure, we could not explain why (d) is grammatical but (c) is not.

It is significant in view of these facts that *he* can modify some actors of a negative sentence; in particular, it is allowed for actors of accusative verbs which have been raised to the higher clause.

(e) kaahore 0 he tangata i patu i te poaka.
neg nom a man tns hit acc the pig

"A man didn't kill the pig."

However, it is not allowed when Raising has failed to apply:

(f) *kaahore i patu 0 he tangata i te poaka.
neg tns hit nom a man acc the pig

These examples imply that *he*-insertion follows Raising, and that the Raising rule turns the actor of the complement into the subject of a higher intransitive verb. This NP must be a subject rather than a topicalized noun, since sentences like (g) are not generally allowed in Maori:

(g) ?he tangata i patu i te poaka.
a man tns hit acc the pig

"..A man killed the pig."

(G) is acceptable only in answer to a question like (h)

(h) ko wai i patu i te poaka?
top who tns hit acc the pig

"Who killed the pig?"

and is not a simple topicalized form.

84. Notice that Topicalization is different for objects of accusative verbs, which require the relative marker ai:

(a) ko Mere i awhiti ai Ø a Pita.
*Ø
top Mere tns help rel nom pers P.

"It is Mere who Pita helped."

And it is not allowed at all for actors of passives, which are preceded by e:

(b) *ko Pita i awhiti - a (ai) Ø a Mere.
top P. tns help pass rel nom pers M.

85. I am indebted to Ross Clark and Tevita 'Ahokovi for the comment on (350), and some of the other examples in this section.

86. This rule has a very small range, since it may not apply to negative sentences:

(a) 'oku 'ikai te u ui Ø e fefinē.
tns not tns I call abs the woman

"I won't call the woman."

But:

(b) *'oku ou 'ikai te ui 0 e fefinē.
ke
tns I not tns call abs the woman

Nor may it apply to higher adjectives like *sai* 'good'.

87. Other negatives in Pukapukan include *yee* 'not', which seems to be inserted after a tense marker and before the negated verb. The fact that *yee* is a negative verb + tense marker in underlying structure is shown by the way it undergoes Clitic Placement:

(a) ko yee a aku patu - *a* 0 te kulii.
tns not pers I hit pass nom the dog

"I won't kill the dog."

(b) ko a aku yee patu - *a* 0 te kulii.
tns pers I not hit pass nom the dog

"I won't kill the dog."

The fact that the clitic appears after *yee* in (a) suggests that some part of it (probably *e*) is actually a tense marker. (B) demonstrates that the negative may also be analyzed as an adverb; *yee* is thus an auxiliary like Samoan *lee*, *le'i*.

88. For instance, Topicalization. I should note that *kiai* can be preceded by the tense marker *e*, i.e.

(a) e kiai na tiaki - na 0 te wenua e Te Malo.
tns not tns lead pass nom the land agt T.M.

"Te Malo didn't lead the country."

but this usage seems to be formal.

89. It also suggests that ergatives cannot be derived by

the Passive rule in Pukapukan.

90. I do not have a terribly clear picture of se-Raising, but it seems to work as follows. Any subject, actor or object preceded by se may be raised to subject position of a higher clause. (The higher clause must naturally be followed by the complementizer 'ona.) Objects of accusative verbs must leave behind a pronominal copy iai, and all other NP's lose their original case markers:

(a) 'ua saa Ø se tangata 'ona ulaula Ø se
*e
tns forb. a man compl smoke abs a

tapa'a.
tobacco

"A man is forbidden to smoke a cigarette."

From:

(b) 'ua saa 'ona ulaula e se tangata Ø se
tns forb. compl smoke erg a man abs a

tapa'a.
tobacco

"It is forbidden for a man to smoke a cigarette."

In (a), Raising applies to the actor of an ergative verb (and the passive suffix -ina may be added without affecting grammaticality). Notice that the ergative e is replaced by Ø.

(c) 'ua saa Ø se tangata 'ona angaleanga 'i
*e
tns forb. a man compl ill-treat acc

lona to'alua.
his wife

"It is forbidden for any
man to mistreat his
wife."

From:

(d) 'ua saa 'ona angaleanga Ø se tangata 'i
tns forb. compl ill-treat nom a man acc

lona to'alua.
his wife

"It is forbidden for any
man to mistreat his wife."

(E) shows the rule applying to objects:

(e) 'ua saa Ø se fafine 'ona angaleanga iai Ø
*Ø
tns forb. nom a woman compl ill-treat rel nom

se tangata.
a man

"Women (in general) should
not be mistreated by men."

From:

(f) 'ua saa 'ona angaleanga Ø se tangata 'i
tns forb. compl ill-treat nom a man acc

se fafine.
a woman

"Women (in general) should
not be mistreated by men."

I do not have examples of se-Raising applying to subjects of intransitives, but it seems reasonable to assume that they undergo the rule like all other NP's.

In general, my data on Raising suggests that the process does not divide clearly between the se- and le-Raising rules; informants give slightly different judgements on sentences, and it may be that semantics as well as syntax is involved. (The data in §6.4 reflects the judgement of the majority of my informants.) However, if we assume that something like *se*-Raising is a rule, it can be formalized as follows.

(TS9) *se*-Raising (optional)

$$\begin{array}{ccccccccc} V & [_{NP} & \underbrace{[_{S}\ Tns\ V\ X} & \left\{\begin{array}{c}\langle \underline{'i} \rangle \\ Case\end{array}\right\} & \underbrace{\underline{se}} & N & \underbrace{Y\ {}_{S}]} & {}_{N}] \\ 1 & 2 & 3 & 4 & 5 & & 6 & 7 \end{array}$$

$$\longrightarrow 1\ \ \underline{\emptyset}\ \ 5\ \ 3\ \ \langle \underline{iai} \rangle\ \ 6$$

It is interesting to compare this and the *le*-Raising rule with the Raising process which we found in Tongan. In Tongan, Raising simply moves the logical subject along with its original case marker; this is *Ø* for subjects of intransitives, but *e* for actors of transitive verbs. In Samoan, this contrast of *e* and *Ø* has been reanalyzed to reflect the contrast of specific/non-specific nouns. Specific nouns which have undergone Raising are preceded by *e*; non-specifics are preceded by *Ø*.

91. *A'oa'o* was originally an accusative verb (Churchward, 1951) but has been reanalyzed as passive/ergative in the speech of my informants. According to them, it would be simply wrong to use the accusative in (373).

92. This is something of an exaggeration, since most

'ona verbs are not subcategorized for personal subjects (and therefore, it is impossible to apply a simple test as in 379-81). However, of the *'ona* verbs which may take personal subjects, all are apparently subcategorized for one NP.

The view that *'ona* introduces sentential subjects is put forth by Marsack (1962). The complementizer is also used for adverbial complements of cause, if the verb of the complement is followed by the particle *lea*.

93. Notice that (D') does not apply to conjoined sentences, and so it is not violated by the only ergative non-clause restricted rule of which I am aware. Dixon has shown that Djirbal, an ergative language, has an ergative rule which refers to NP's of successive (conjoined) clauses. This rule of Topic Chains apparently does not go down into complex sentences, and thus does not constitute a counter-example to (D').

Bibliography

Anderson, Stephen R. MS, 1968. On grammatical relations in two Polynesian languages.

Beaglehole, Ernest, and Pearl Beaglehole. MS, 1938. Myths, stories and chants from Pukapuka. Honolulu: Bernice Pauahi Bishop Museum.

Biggs, Bruce. 1969. Let's learn Maori. Wellington, N.Z.: A.H. and A.W. Reed.

____________. MS, 1971. The Passive Voice in Polynesian Languages.

Chapin, Paul G. 1970. Samoan pronominalization. Lg. 46.366-78.

Chomsky, Noam. 1965. Aspects of the theory of syntax. Cambridge, Mass.: MIT Press.

_____________. 1970. Remarks on nominalization. Readings in English transformational grammar, ed. by R. Jacobs and P. Rosenbaum, 184-221. Waltham, Mass.: Ginn and Co.

Chung, Sandra. MS, 1970. Negatives as verbs in Polynesian. Cambridge, Mass.: Harvard University senior honors thesis.

Churchward, C. Maxwell. 1953. Tongan grammar. London: Oxford University Press.

____________________. 1959. Tongan dictionary. London: Oxford University Press.

Churchward, Spencer. 1928. On the origin of the Polynesian passive. J.Polynes.Soc. 37.300-5.

_________________. 1951. A Samoan grammar. Melbourne: Spectator Publishing Co. Pty. Ltd.

Clark, Ross. MS, 1971. On the history of negation in Polynesian languages.

Fillmore, C.J. 1968. The case for case. Universals in linguistic theory, ed. by E. Bach and R. Harms, 1-88. New York: Holt, Rinehart and Winston.

Frisbie, Robert Dean. 1930. The book of Puka-Puka. London: John Murray.

Grezel, le P. 1878. Dictionnaire Futunien-Français avec notes grammaticales. Paris.

Grinder, John. 1969. Conjunct splitting in Samoan. Linguistic notes from La Jolla, no. 2, 46-79.

Hale, Kenneth. MS, 1968. The passive and ergative in language change: the Australian case.

______________. 1968a. Review of A profile generative grammar of Maori, by Patrick W. Hohepa. J.Polynes. Soc. 77.83-99.

Hohepa, Patrick W. 1967. A profile generative grammar of Maori (Indiana University Publications in Anthropology and Linguistics, Memoir 20).

__________________. 1969. The accusative-to-ergative drift in Polynesian languages. J.Polynes.Soc. 78.295-329.

__________________. MS, 1970. Not in English and kore and eehara in Maori.

Johnson, Alan, and Lillian Harmon. 1970. Let's speak Samoan. Auckland, N.Z.: Church of Jesus Christ of Latter-day Saints.

Kiparsky, Paul. 1968. Linguistic universals and linguistic change. Universals in linguistic theory, ed. by E. Bach and R. Harms, 171-202. New York: Holt, Rinehart and Winston.

_______________. 1971. Historical linguistics. A survey of linguistic science, ed. by William O. Dingwall, 576-642. University of Maryland.

Language Staff of Peace Corps Polynesia, Samoa II. 1968. A Samoan language manual.

Lynch, John. 1969. Voice and case in Tongan. Working papers in linguistics, no. 6, 71-85. Honolulu: University of Hawaii.

Marsack, C.C. 1962. Teach yourself Samoan. London: The English Universities Press.

McEwen, J.M. 1970. Niue dictionary. Wellington, N.Z.:

Department of Maori and Island Affairs.

Milner, G.B. 1962. Active, Passive or perfective in Samoan: a fresh appraisal of the problem. J.Polynes.Soc. 71.151-61.

___________. 1966. Samoan dictionary. London: Oxford University Press.

Neffgen, H. 1918. Grammar and vocabulary of the Samoan language, transl. from the German by Arnold B. Stock. London: Kegan Paul, Trench, Trubner and Co., Ltd.

Pawley, Andrew K. 1966. Polynesian languages: a subgrouping based on shared innovations in morphology. J.Polynes. Soc. 75.39-64.

________________. 1966a. Samoan phrase structure: the morphology-syntax of a western Polynesian language. Anthropological Linguistics 8, no. 5.

________________. 1967. The relationships of Polynesian Outlier languages. J.Polynes.Soc. 76.259-296.

Pratt, George. 1862. Samoan dictionary. Samoa: London Missionary Society's Press.

Senn, Alfred. 1966. Handbuch der Litauischen Sprache, Band I: Grammatik. Heidelberg: Carl Winter, Universitatsverlag.

Sharples, Peter. 1968. Sikaiana syntax. Auckland, N.Z.: University of Auckland masters thesis.

Shumway, Eric B. 1971. Intensive course in Tongan. Honolulu: University of Hawaii Press.

Stuebel, O. MS, n.d. Selections from Samoan texts, 1884-94. Library of American Samoa.

Tregear, Edward. 1891. The Maori-Polynesian comparative dictionary. Wellington, N.Z. (reprinted by Oosterhout N.B., The Netherlands, 1969).

Williams, Herbert W. 1928. Some observations on Polynesian verbs. J.Polynes.Soc. 37.306-17.

__________________. 1971. A dictionary of the Maori language, 7th edition. Wellington, N.Z.: A.R. Shearer, Government Printer.

Williams, W.L., and H.W. Williams. 1965. First lessons in Maori (13th edition, revised by W.T. Ngata). Christchurch, N.Z.: Whitcombe and Tombs.

Notes on the Formation of Complex Sentences in Tunisian Arabic with Emphasis on the Relative Clause

Colette G. Craig

This paper is a presentation of the formation of complex sentences in Tunisian Arabic.[1] I have taken one type of embedding, the relative clause construction, to be the main focus of investigation. I will show how the relative clause in Arabic functions as an adjectival qualifier and how the facts of relativization produce no change in the constituent structure of the embedded sentences. This means that an embedded sentence in Arabic could be pulled out of its subordinate position and stand by itself as an independent sentence. I will then look briefly at another type of embedding, the complement sentence, and show that the same principle applies there as well. Finally, I will consider implications of the assumption that all clauses in Arabic are full, potentially independent clauses simply juxtaposed and concatenated: I will specifically argue that this will account for the structure of the topicalized sentences and the ungrammaticality of reduced coordinate structure in the language.

1-0 The relative clause in Tunisian, like that of English, is a clause modifying an NP in which the clause contains a NP coreferential to the modified NP.

The study of relative clauses is a double one: the first task is to determine the formal means by which the language realizes this relation (i.e., what arrangements of constituents and insertion of particles are called for in the language under investigation); the second is to determine the exact syntactical function of this type of clause. As Benveniste pointed out:

> Le problème est de déceler cette fonction. On peut y parvenir en observant que souvent la phrase-relative a, dans le système linguistique considéré, les mêmes marques formelles qu'un autre syntagme, de dénomination toute différente, (...) Guidée par cette analogie formelle, l'interprétation de la phrase relative devient possible en terme de fonction. C'est un rapport *interne* que nous nous proposons de mettre en lumière.[2]
> (emphasis Benveniste's)

1-1 Let us start by analyzing the formal structure of a relative clause. Consider the following sentences:

1- il ra.žil illi (huwwa) mša. li-il blad
the man REL. he went to the town
'the man who went to town'

2- il ṭawla.t illi ža.b-hum
the tables REL brought-them F. pl.
F. pl.
'the tables that he brought'

3- il ṭawla *illi* ka.n fuḳ-ha il kta.b
the table *REL* was on-her the book
F. sg. F. sg.
'the table on which the book was'

In sentences 1-3 the antecedents of the relative clauses are (+definite). They are all preceded by the definite article *il*.[3] In this (+definite) environment the relative clause is introduced by a particle, *illi*. *Illi* is not a relative pronoun because the coreferential NPs inside the embedded clauses are still present as pronouns. *Illi* does not inflect for gender or for number. It is a non-anaphoric particle which is inserted at the head of the relative clause only when the antecedent is (+definite) as the next sentences will show:

4-a huwwa waḥad ma. y-istana-š
he(is) one not he-waits -neg
'he is one who does not wait'

-b huwwa *il* waḥad *illi* ma. y-istana-š
he the one *REL* not he waits-neg
'he is the one who does not wait'

5-a ama šuf-t mra waḥda raḳda
I saw-I woman one is sleeping
'I saw one woman who was sleeping'

-b ana šuf-t il mra illi raḳda
I saw-I the woman REL sleeping
'I saw the woman who was sleeping'

The illi insertion rule looks like:

SD:	X	NP_1(	DET	N	)	Y	NP_1	Z
	1		2	3		4	5	6

SC insert illi between 3 and 4

1-2 Let us now turn to what happens inside the relative clause. As mentioned above, the coreferential NP is not moved to the head of the sentence, nor does any kind of relative particle get attached to it. Rather, it is found pronominalized by the same Pronominalization transformation that is otherwise independently motivated. The same transformation will take care of 6 through 11:

6- Našžl.b ša.f bint il Nažl.b
Nažl.b ša.f bint-u
'Nažl.b saw his daughter'

7- il dar ka.n ba.b il dar kbi.r[4]
il dar illi ka.n ba.b-ha kbi.r
'the house the door of which is big'

8- ana šuf-t Nažl.b w-ana ḳal-t-li Nažl.b
ana šuf-t Nažl.b w-ana ḳal-t-li-u[5]
'I saw Nažl.b and I talked to him'

9- il ra.žil ana n-ḳal-li- il ra.žil
il ra.žil illi ana n-ḳal-li-u
'the man that I talked to (him)'

10- il ḳsar inta t-uskun fi-il ḳsar
il ḳsar illi inta t-uskun fi-h
'the castle that you live in (it)'

11- il ša.b il bulisi ka.n y-ižri ura il ša-b
il ša.b illi il bulisi ka.n y-ižri ura-hum
'the young men that the policeman is running after (them)'

The Pronominalization transformation will turn obligatorily the second of two coreferential NPs into the appropriate pronoun.[6] It can be written as:

SD:	X	NP_i	Y	NP_i	Z	
	1	2	3	4	5	
SC	1	2	3	Pro+4	5	where 2 and 4 are coreferential

1-3 The so-called 'returning pronouns' of the relative clauses (pronoms de rappel) do not always show on the surface structure. The subject pronoun is optionally deleted, as it is in independent sentences, when it is followed by an inflected verb carrying a person marker which is not ambiguous (therefore not with second and third person singular feminine in the imperfect, nor with first or second person singular in the perfect):

t-uskun	you/she live(s)
šuf-t	I/you saw

It can always be deleted when the context available is sufficient to disambiguate the verb form:

(huwwa)	ša.f	he saw; always optional
ana inta	šuf-t	I/you saw; optional if sufficient context provided

The returning subject pronoun of a nominal predicate is optional in its independent form (huwwa, hiyya, ...) but remains obligatorily in a compound form (with the negative for instance); in an independent clause the subject pronoun of a nominal predicate is obligatory, in the view of the absence of a copula:

12-a	huwwa	baba	'he is my father'
-b	hiyya	amira	'she is a princess'

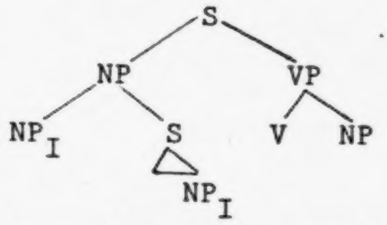
13-a	huwwa	dima muriḍ	'he is always sick'
-b	ma-hu-š (huwwa)	ɣniɣ	'he is not rich'

14- il ra.žil illi (huwwa) ža. li-il blad
the man REL he went to the town
'the man who went to town'

15- il ra.žil illi hna (huwwa) baba
the man REL here he(is) my father
'the man who is here is my father'

16- il ra.žil illi ma-hu-š ɣni
the man REL neg-he-neg.
(huwwa)
'the man who is not rich'

The optionality of the object returning pronoun does not follow such a clearcut pattern. The pronoun is obligatory when the antecedent is a definite noun fully specified, for example by another relative clause:



17- t-'arfi-š il ra.žil illi ɣni yasir illi ana šuf-t-u
do you know the man REL rich very REL I saw - him
'do you know the man that I saw, who is very rich'

18- t- 'arif-š il ra.žil il smi.n u illi 'andu šašiya
do you know the man the fat and REL he has a hat

ḥamra w-illi ana ḳabil-t-u
red and REL I met -him
'do you know the fat man that I met who has a red hat (sheshya)'

It is also obligatory when the antecedent is a (-definite) NP:

19- ra.žil ana šuft-u
a man I saw-him
'a man that I saw'

20- mra (huwwa) ša.f-ha
a woman he saw-her
'a woman that he saw'

It becomes optional when the relative clause does not have any overt antecedent:

21- ana šuf-t illi Nažī.b ša.f-(u)
I saw REL Nažī.b saw-him
'I saw who Nažī.b saw'

22- nažžim-š inta t-ṭalla illi (huwwa) ḍ̌rab-(u) il baraḥ
is it possible you guess REL he hit-him yesterday
'can you guess who he hit yesterday'

The remaining optional object pronouns appear in sentences that seem to have in common a limited information content, either that the relative clause is only loosely attached to the antecedent (as in 23a), or that the content of the relative clause is highly predictable (as in 24):

23-a il mra illi huwwa ša.f
'the woman that he saw'

-b il mra illi huwwa ša.f-ha
'the woman that he saw'

According to the informant's intuition, ša.f in 23-a is a verb with no object which describes an action that is not very tightly connected to the existence of the woman. A suggestive translation of that sentence would be: "there was the woman, and there was he seeing, and he saw her"; while 23-b emphasizes that what he did see was precisely the woman.

24- il banana illi ana kli.t
the banana REL I ate
'the banana that I ate'

compared to:

25- il ṭawla illi ana kli.t-ha
the table REL I ate-it(her)
'the table that I ate'

It is difficult to set any kind of formal device to account for these optional deletions of 'predictable' objects like a feature [+predictable]. This deletion remains a marginal fact which should probably be dealt with at the speech performance level.[7]

In the cases where the verb of the relative clause is transitive, the possible combinations of optionality could yield ambiguous sentences when both object and subject are third person singular of the same gender. Consider:

26- il ra.žil illi (huwwa) ša.f-u
opt. oblig.
the man REL he saw- him
'the man that saw him'

27- il ra.žil illi (huwwa) ša.f-(u)
opt. opt.
the man REL he saw- him
'the man that he saw'

The two possible ambiguous sentences that could be produced are:

28-a il ra.žil illi huwwa ša.f-u

-b il ra.žil illi ša.f-u

The distribution in actual speech seems to be:

-c il ra.žil illi huwwa ša.f
for 'the man that he saw'

-d il ra.žil illi ša.f-u
for 'the man that saw him'

even though: il ra.žil illi huwwa ša.f-u is used when the context is there to disambiguate it.

In possessive constructions and in prepositional phrases the returning pronouns are always obligatory:

29- il dar illi ka.n ba.b-ha kbi.r
the house REL was door-her big
'the house the door of which was big'

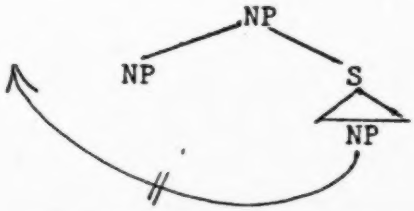
* il dar illi ka.n ba.b __ kbi.r
the house REL was door __ big

30- haďa il fard illi ana ḳtal-t bi-h il kalb
this the gun REL I killed with-it the dog
'this is the gun that I killed the dog with'

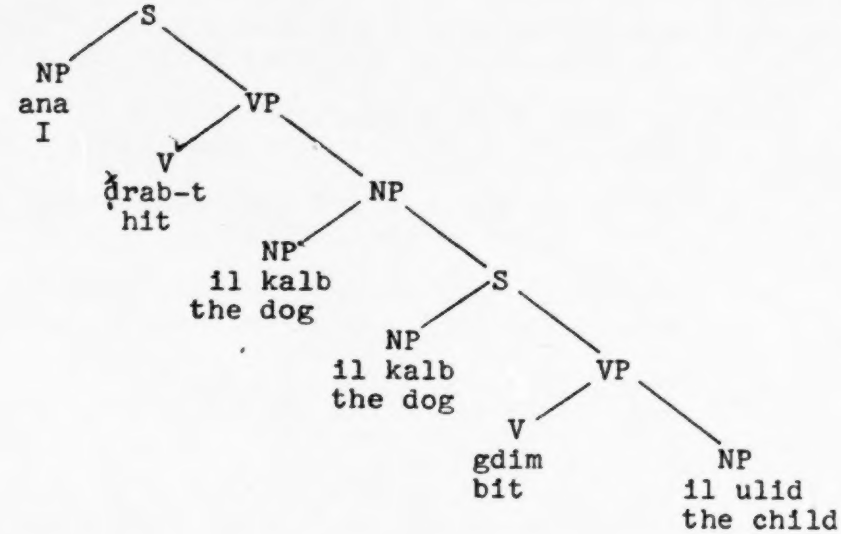
* haḍ̌a il fard illi ana ḳtal-t bi il kalb

In conclusion, the elements that can be deleted in a relative clause in Arabic are the same elements that can be deleted or absent in an independent clause, like the subject, or the object of a transitive verb that can be used without it.[8] But the language will not allow the deletion of any necessary element of a constituent structure that would be unrecoverable in an independent sentence as would be the case for the noun of a prepositional phrase.

1-4 One of the consequences of the fact that in Arabic the relativization process does not require any movement rule is that the Complex NP constraint that Ross established for English does not obtain in the language.[9] Ross' constraint is that no NP of an S which is dominated by another NP can be pulled out of that lower S:



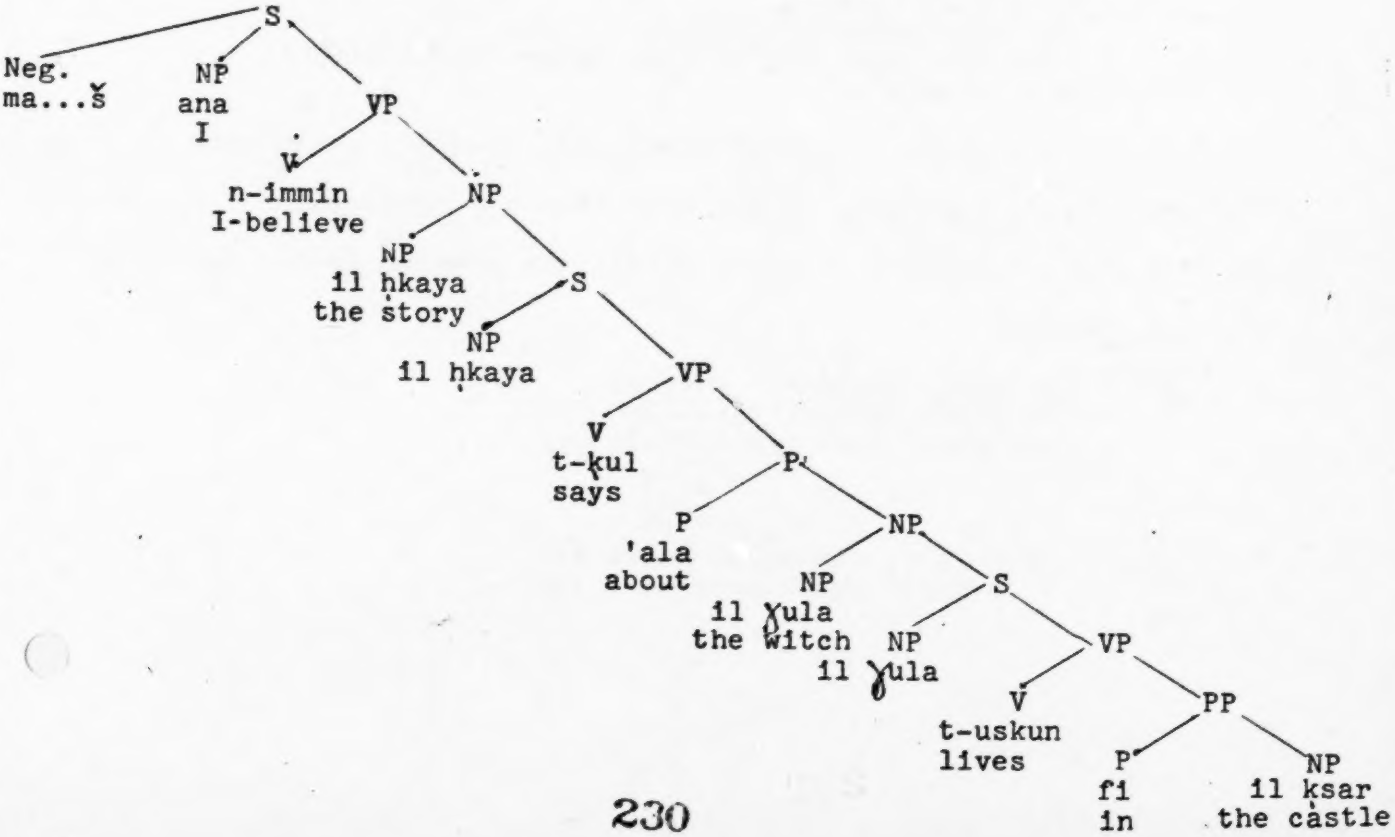
But with sentence 31 to which the following tree corresponds, sentence 32 can be produced:


S
NP
ana
I
VP
V
ḍrab-t
hit
NP
NP
il kalb
the dog
S
NP
il kalb
the dog
VP
V
gdim
bit
NP
il ulid
the child

31- ana ḍrab-t il kalb illi gdim il ulid
I hit-I the dog REL bit the child
'I hit the dog who bit the child'

32- il ulid illi ana ḍrab-t il kalb illi gidm-u ka.n yipki
the child REL I hit-I the dog REL bit-him was he-cries
'the child that I hit the dog that bit him was crying'

In the same way 34 is derived from 33 reproduced on the tree:


S
Neg.
ma...š
NP
ana
I
VP
V
n-immin
I-believe
NP
NP
il ḥkaya
the story
S
NP
il ḥkaya
VP
V
t-ḳul
says
P
P
'ala
about
NP
NP
il ɣula
the witch
S
NP
il ɣula
VP
V
t-uskun
lives
PP
P
fi
in
NP
il ksar
the castle

33- ana ma. n-immin-š bi il ḥkaya illi t-ḳul 'ala
I neg. believe-neg. the story REL says about

il ɣula illi t-uskun fi-il ḳsar
the witch REL lives in the castle

'I do not believe the story that tells about the witch
that lives in the castle'

34- il ḳsar illi ana ma. n-immin-š bi il ḥkaya illi
the castle REL I negI-believe.-neg part, the story REL

t-ḳul 'ala il ɣula illi t-uskun fi-h ka.n fi ras il žbal
says about the witch REL she-lives in it was on top(of)
the mountain

'the castle that I do not believe the story that tells
about the witch that lives in it, was on top of the
mountain'

In Arabic an NP can be pulled out of an adverbial phrase too:

35- Naži.b ža. b'ad ma. kammil il xidmat
Naži.b came after that he finished the work
'Naži.b came after he finished his work'

36- il xidma illi Naži.b ža. b'ad ma. kammil-ha
the work REL Naži.b came after that finished-it (her)

ka.nit mit'aba
was exhausting
'the work that Naži.b came after he finished it
was exhausting'

Tunisian also has relativization with wh-words which move to the head of the sentence, taking with them the preposition. The wh-word which functions fully as a relative pronoun in our data is wi.n "where":

37- il dar fi-wi.n t-uskun
the house in-where you-live
'the house that you live in'

38- ana n-'araf min-wi.n huwwa ža.
I I-know from-where he came

39- ana n-'araf li-wi.n huwwa mša.
I I-know to-where he went

It is also a question word:

40- fi-wi.n illi ḍ̌rab-ni txabba?
in-where REL hit-me hid
'Where did the one who hit me hide?

41- Wi.n t-uskun?
'Where do you-live?'

As expected, these relativized sentences with anaphoric wh-words do not have returning pronouns.[10] In a predictable manner too, the constraint applies to the sentences with wh-words while it did not to sentences with illi:

42- il banke illi Naži.b ka.n farḥan b'ad ma. mša.-li-ha
the bank REL Naži.b was happy after that went to it
(her)

ka.n-it fi-il blad
was(F.) in the town
'the bank that Naži.b was happy after he went to it,
was in the town'

43-* il banka fi-wi.n Naži.b ka.n farḥan b'ad ma. mša.
the bank in where Naži.b was happy after (that) he went

ka.n-it fi-il blad
was (F.) in the town.

To recapitulate this presentation of the relative clause in Arabic: The relative clause is transformationally derived by two rules, that of illi insertion, which is dependent on the presence of a definite antecedent, and that of pronominalization, which is followed by rules of optional deletions of some subjects and objects.

2-0 Now that the formal structure of the relative clause is defined, it remains to see if there exists another formal construct of the language which, by identity or similarity with it, would reveal the internal function of the relative clause in Arabic.

2-1 A striking parallelism with the structure of the adjectival NP imposes itself. By adjectival NP I mean a node NP which dominates a noun modified by an adjective. In adjectival NP, the adjective not only agrees in gender and number with the head-noun but it also agrees in definiteness:

44-a ţawla kbi.r-a a big table (F. sg.)
-b rža.l sma.n fat men (M. pl.)

45-a il ţawla il kbi.r-a the big table
-b il rźa.l il sma.n the fat men

This construction of the adjectival NP includes derived adjectives corresponding to the English participle -ing:

46- il ulid il ra.k̩id from il ulid y-irķid
the child the sleeping
'the sleeping child' 'the child sleeps'

47- il ulid il kaθab from il ulid y-ikðib
the child the lying
'the lying child' 'the child lies'

48- il ra.žil il sarak̩ from il ra.žil y-israk̩
the man the stealing
'the stealing man' 'the man steals

These particular constructions can also receive illi as in: 'il ulid illi ra.k̩id'. At the same time others cannot. Whether this restriction is a reflection of phonological or semantic rules remains to be investigated.

2-2 I will include as part of the same construction the possessive construction of the type 'the king's daughter' which is rendered in Tunisian as NP il NP:

49-a bint il malak 'the king's daughter'
*-b il bint il malak

50-a dabbuzit il šra.b 'the bottle of wine'
*-b il dabbuzit il śra.b

The peculiarity of this Arabic construction is the absence of the definite article of the head-noun even though it is (+definite). It also shows constraints on the appearance of adjectives. The (+definite) 'bint' in 49-a cannot have an adjectival form immediately following it:

51-a * bint il źmila il malak
daughter the beautiful the king
'the king's beautiful daughter'

-b bint il malak, (hiyya) źmila
'the king's daughter, she is beautiful'

while the noun with an article can.

52- bint il malak il ḳasi
daughter the king the fearful
'the fearful king's daughter'

2-3 Notice here that a noun cannot take more than one adjective at a time:

53-a il dar il ḥamra the red house
-b il dar il kbi.ra the big house
-c * il dar il kbi.ra il ḥamra
-d il dar il kbi.ra, lun-ha aḥmar 'the big house its (her) color is red'

54-a il kalb il kbi.r 'the big dog'
-b il kalb il muriḍ 'the sick dog'
-c * il kalb il kbi.r il muriḍ
-d il kalb il kbi.r w-illi huwwa muriḍ ma. habbi-š
the dog the big and REL he(is) sick neg. wants-neg.

y-akul
eats
'the big sick dog does not want to eat'

These two facts -- the presence of a definite article in front of the adjective when the head-noun is definite, and the constraint on the number of adjectives -- should be accounted for in some uniform way. There is a simple way out: consider the construction 'il NP il Adj.' as a relative clause, where the second article functions as a relative marker. This would explain why a single noun will not be modified by more than one adjective. The way 'the big fat ugly ... dog' is produced in English is by coordintion of sentences and then reduction of these coordinated structures. We shall see later that conjunction reduction is not free in Arabic as it is in English. Sentences like 54-d are possible as coordinated sentences but are not very natural Arabic sentences. The preference of the informant was for two separate statements; for example, in 54 that the dog was big was not related to its being sick.

2-4 Compare the following pairs of sentences:

55-a il ulid illi huwwa ra.ķid 'the child who is sleeping'
-b il ulid il ra.ķid 'the sleeping child'
-c il ulid illi huwwa dima farḩan 'the child who is always happy'
-d * il ulid il dima farḩan

56-a ulid huwwa ra.ķid 'a child who is sleeping'
-b ulid ra.ķid 'a sleeping child'
-c ulid huwwa dima farḩan 'a child who is always happy'
-d * ulid dima farḩan

Relative clauses of the type: 'il ra.žil illi huwwa ɣni' share with the adjectival construction 'il ra.žil il ɣni' the formal angle-bracket device which says: if $_a\langle\ \rangle_a$ then $_b\langle\ \rangle_b$. It remains to account for the difference between the second terms of the angle bracket, il and illi.

The relationship of il ... illi ... to il ... il ... is the one of a full relative clause to a reduced one.

Look at:

57-a il ra.žil il farḩan 'the happy man'
-b * il ra.žil il $\left.\begin{array}{l}\text{ma}\\ \text{la}\end{array}\right\}$ farḩan 'the not happy man'
-c * il ra.žil il dima farḩan 'the always happy man'

58-a il dar il asɣar 'the small house'
-b * il dar il asɣar min dar-ik 'the house smaller than your house'
-c il dar illi asɣar min dar-ik

It would probably be more appropriate to consider the full relative clause as an expansion of the adjectival NP rather than the adjectival NP to be the reduced form of the full relative clause. The adjectival NP with simple article il will expand any time another element (negative particle, adverb, comparative particle) modifies the adjective.[12]

Languages fall generally into one or the other of the following types of relativization: i) where the relative marker is a wh-word; and ii) where the relative marker is a definite article. Tunisian Arabic has both types, with wh-words like wi.n which behaves freely as relative and question marker, and with illi which makes of the relative clause a paratactic construction. The internal function of the relative clause has been described as one of an expanded adjectival construction in apposition to the head-NP.

3-0 In this last part of the paper I will show several cases where the same returning pronouns introduced in the discussion on relative clauses appear. I postulate that their presence reflects a general Arabic constraint on the constituency of a clause, be it independent or subordinate.

3-1 Here are a few examples of complement sentences:

59- ana n-ḥab baš n-imši
I I-want aspect I-go
'I want to go'

60- ana n-ḩab-ik baš t-imši
I I-want-you asp. you-go
'I want you to go'

61- ana ţlab-t baš n-iftah il ba.b
I asked asp. I-open the door
(incompletive)
'I asked to open the door'

62- 'Ali ža. baš i-žib taffaḩa
Ali came asp. he-brings apples
(incompletive)
'Ali came to bring apples'

These complement sentences are not introduced by any complementizer. They are simply juxtaposed to the main verb. Baš is an aspect marker which indicates a future with a notion of probability in independent sentences, as in sentences 63-64:[13]

63- ana baš n-imši li-il blad
I will I-go to the town
'I will go to town'

64- baš i-kun ɣudwa
will it-happens tomorrow
'it will happen tomorrow'

65- ana baš n-imši baš n-šuf il amir
I will I-go will I-see the prince
'I will go to see the prince'

66- huwwa ža. (baš) i-šuf-ni
he came will he-sees-me
'He came to see me'

It also marks, as in 65, 66, one of the two possible aspects of a subordinate verb. It has not been possible to determine in which environment it has to occur. In the available data it seems to alternate freely with the incomplete form of the verb. In the set of examples given above (59 to 62, 65 to 66) the complement sentences express a goal or a desire.

In the cases where the subject of these complement sentences is coreferential with either the subject, or the object, of the main verb, Arabic does not have an Equi NP deletion rule. The subordinate verb agrees in person with its subject, which has been pronominalized. This same subject is nevertheless optionally deleted later whenever it is not emphasized.

The pronominalized complement NP which shows in the main sentence could be the result of subject raising. If it is subject raising, the rule will have to be a copying rule rather than a movement rule, and will not be restricted to the position of direct object of the main verb:

67- ana n-ḥab-ha baš t-kun haḍra
I like-her asp. she is ready
'I like her to be ready'

68- huwwa kla-l-i ḥuta baš n-akul-ha
he fried-for-me a fish asp. I-eat-it (her)
'he fried for me a fish to eat'

More likely the pronoun in the main sentence is the anaphoric form of an NP which was already in the deep structure of that sentence, to which a complement sentence has been juxtaposed.[14]

3-2 The next set of examples are complement sentences on the other hand introduced by a complementizer inn-:

69- huwwa faḳḳar-ni inn-i ku.nt kalb
he reminded-me comp.-I was-I a dog
'he reminded me that I was a dog'

70- huwwa faḳḳar-ni inn-i baš n-imši
he reminded-me comp.-I asp. I-go
'he reminded me that I should go'

71- 'ali ḥlaf inn-i baš n-mu.t
'Ali swore that-I {will} die-I'
{would}

72- ana n-'araf inn-u il šra.b ka.n fuḳ il ṭawla
I I-know comp.-it the wine was on the table
'I know that the wine was on the table'

73- huwwa ɣabar-ni (inn-u) baš i-ži
he announced-me (comp.-he) asp. he-come
'he announced me that he would come'

All the examples found in the data and presented here are verbs of saying: say, remind, swear, announce, (and know) that... The aspect of the subordinate verb is not restricted. It can be completive (69, 72). The aspect word baš here takes a definite conditional value. It is not the purpose of this paper to analyze in detail the structure of the complement sentence introduced by inn-, but let us say that inn- might be at some point analyzed as another appearance of illi, as the following pair of sentences indicate:

74-a bu-na faḳḳar-na illi ahna kun-na ska.ra
father-our reminded-us comp. we were-we drunk
'our father reminded us that we were drunk'

-b huwwa faḳḳar-ni inn-a kun-na ska.ra

75-a huwwa faḳḳar-ni illi ana ku.n-t kalb
he reminded-me comp. I was-I a dog
'he reminded me that I was a dog'

-b huwwa faḳḳar-ni inn-i ku.n-t kalb

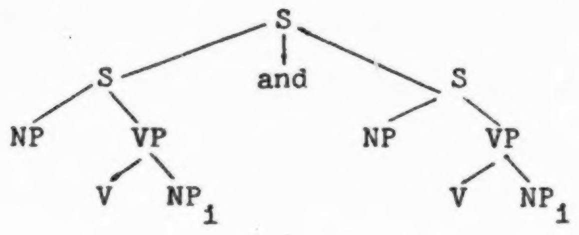
This complementizer inn- is completely optional and is a characteristic feature of the speech of the educated speakers.[15]

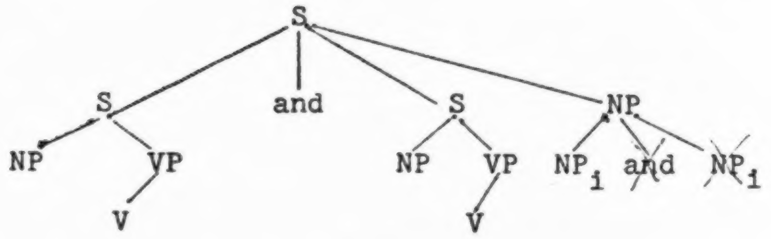
The clitic pronoun that is suffixed to the complementizer must be an extra pronoun as shown in sentences which already have a third person subject:

76- ana ḳul-t inn-u 'ali baš i-mu.t
I said-I comp.-he Ali asp. he-dies
'I said that Ali would die'

This would be a new area of syntax to explore.

3-3 Coordination in Tunisian Arabic is done by a conjunction u/w which precedes the first element of the second sentence. English has a rule of coordinate structures reduction which has the effect of rearranging the structure of the sentence by pulling out the coreferential NPs, and then deleting one of them:





This operation is impossible in Tunisian:

77-a Dalila žabi-t il hwa.ž u Samira twa.-t il hwa.ž
'Dalila brought the laundry and Samira folded the laundry'

-b * Dalila žabi.t u Samira twa.t il hwa.ž
-c * Dalila žabi.t-hum u Samira twa.t il hwa.ž

Arabic does not allow for the restructuring of a sentence which would produce an empty slot in the tree. Neither does it have backward pronominalization in independent sentences that would remedy it.[16]

3-4 Topicalization

There are two ways that an NP can become the focus of the sentence. Either it receives a mark (as is the case in Japanese, a case-marker) or it moves, to the head of the sentence most often, from its original, well defined place. In Arabic the topicalized NP appears at the head of the sentence. If it is part of a prepositional phrase, the NP alone is found in initial position with the preposition still behind in its original place. Generally Topicalization is considered to be a movement rule that creates a rearrangement of the constituents. It is probably in Tunisian just a copying rule which duplicates the topicalized NP at the head of the sentence and then turns the second NP into a pronoun by the general Pronominalization rule:

78-a i-ži li Samira žwa.b kul yum
'arrives to Samira a letter every day'
-b (Top copying) Samira, i-ži-li Samira žwa.b kul yum
-c (Pron.) Samira, i-ži-li-ha žwa.b kul yum.
'to Samira, a letter arrives every day'

79- il ḫuta il kbi.ra xalli-na-ha li Naži.b
the fish the big we left- it(her) to Naži.b
'the big fish we left for Naži.b'

80- 'ali mart-u ma. t-hayyir-u-š
Ali wife-his not she bothers-him-neg.
'as for Ali his wife does not bother him'

We have shown here how the language resists the breaking up of constituents and the deletion of elements in general, using paratactic constructions as a major way of forming complex sentences. The next step for such a study of Tunisian syntax would be to check what happens with other movement rules as cleft and pseudo-cleft transformations. It would help see how general are the processes presented in a tentative way in this paper.

Notes

1- The data was gathered during the sessions of a field-method course. 'ali Draouil, the informant, was a native speaker from Bizerte, educated in classical Arabic, and spoke the city variant of Tunisian. He occasionally gave classical forms in translation - ellicitation when colloquial Tunisian lacked the form. The only classical form presented in the paper will be identified as such (complementizer *inn-*).

2- E. Benveniste, "La phrase relative, problème de syntax générale" in *Problèmes de linguistique génerale*. Paris, 1966, p. 208.

3- The Tunisian definite article *il* has a broader range of use than the corresponding English *the*: *il* is used generically as in:

- *il* sahbu	a friend (is a man who...)
- t-ɣa.ni kif *il* bulbul	she sings like a bird
- il xanfira ķad *il* ķuffa	the nostrils as big as a bag (couffin)

and with abstract nouns as in:

- ana ma. šuf-t illa *il* ḥamra wa *il* mṣi.ba
  'I did not see but catastrophy and tragedy'

4- All through the paper, a strict phonemic transcription is used. The -l of the article in Arabic assimilates to the first consonant of its head-noun that would be (+anterior) il dar → id-dar
but il ķsar → il ķsar

5- The third person masculin singular clitic pronoun has two allomorphs -u follows a consonant;
-h, whose phonetic realization remains to be established but which is felt consciously to be there by educated people

(another example of the psychological reality of a phoneme), follows a vowel. Nevertheless, Arabic has an alternation

+ li-h+ where li is a free preposition
-l-u + where li is incorporated in the verb and looses its vowel.

6- Even though the coreferential NP is not moved up to the head of the relative clause there is a certain re-ordering inside the relative clause:

il kta.b ka.n fuķ il ţawla
il ţawla illi ka.n fuķ-ha il kta.b
'the table on which the book was'

7- The data given by the informant at intervals of several sessions is not always very consistent. The times when the deletion appears are nevertheless characteristically in sentences with no context where the verb of the relativized clause can be taken for a form with no object.

8- In a similar way, while conjunction reduction is impossible, as will be shown in the last part of the paper, one can obtain a possible reading of:

-Dalila ķašrit u Samira xaslit il batata
'Dalila peeled and Samira cut the potatoes'

if 'to peel' is taken as a verb with no object meaning that Dalila was peeling but not especially the potatoes.

9- The constraint applies to movement rules in general and includes Topicalization, cleft-sentences, questions. Only relativization was checked with the informant. It is assumed that no constraint on the other exists in Arabic either but data should be carefully collected.

10- Tunisian Arabic has another wh-word. šku.n 'who' and aš 'what' are primarily question words which, too, are moved to the head of the sentence with their preposition:

- šku.n ža. il baraḥ
  'who came yesterday?'
- li šku.n ana baš n-'atti il ḳlam mt'a-i
  'to whom will I give the pencil of mine?'
- m'a šku.n-ha huwwa ža
  'with whom (Fem.) did he go?'
- šku.n-ha illi ḍarbit il ba.b
  'Who is she that knocked on the door?'

They can also be relative pronouns . They are limited to the cases that could be rendered by an indirect question sentence:

- huwwa ḳall-i li-šku.n sallif il flus
  'he told me to whom he lent the money'
- ana n- 'araf šku.n ža. il baraḥ
  'I know who came yesterday'
- aš 'and-ik il yum
  'what have you got today?'
- inšd-u bi-aš ḍarb-u
  'ask him with what they hit him'

11- The informant also gave the example:

- bint il malak il žmila mša li-il blad
  daughter the king the beautiful (Fem.) went to the town
  'the king's beautiful daughter went to town'

and he added "but what does the fact that she is beautiful have to do with her going to town?" It is clear for him that her being beautiful is one statement expressed by one sentence, and her going to town another one that he was reluctant to express in the same sentence.

12- *il* and *illi* could be treated just as allomorphs. *il* would then be said to appear in the environment of adjectives and *illi* in all other environments. But this alone would not account for the assymetry of the constructions:

a- il + adj.

b- $\text{illi} + \left\{\begin{matrix}\text{adv.}\\ \text{neg.}\end{matrix}\right\} + \text{verb} = \text{adj.}$

for the fact that *b* could be an independent sentence but *a* could not and therefore that the status of *a* and *b* are not similar.

13- *baš*, as a future marker, is an innovation of colloquial Arabic. The classical language has a form *amši* = 'I will go'. It could be that *baš* is really more than an aspect, maybe a conjunction even as it happens to be in other colloquials. In rural Tunisian of Takrouna studied by W. Marçais, it is said to be a conjunction of finality corresponding to the wh-word, a compound of the preposition *bi* *aš* meaning 'with what, through what'... In Moroccan Arabic it has the same value of conjunction; as a matter of fact, it is a wh-word.

le mkohla *b-aš* ķtel-t s-sbeʿ
'the rifle *that* I killed the lion *with*'

Interestingly enough, the wh-word corresponding to the Tunisian *wi.n* is this *-aš* (which in Tunisian is a question element):

s-sfina f-aš kanu rakbin
'the ship *that* they are riding *in*'

(In *Reference Grammar of Moroccan Arabic*, p. 164, by R. Harrell)

14- More study is needed to determine the exact nature of *baš*. If it is the same *baš* that is found in independent sentences and embedded complement ones, one has to classify it as an aspect marker. But *baš* in complement sentences seems to function sometimes as a complementizer, for instance if it precedes the person marker:

- ana n-ḥab-ha baš *hiyya* t-ku.n ḥaḏ̣ra
  I like-her comp. she is ready
  'I like her to be ready'

15- I am not sure that *inn*- is, as the informant suggested, to be identified as a form of *illi*. It could be a form of the real conjunction like 'an+clitic that shows up in Mauritania. It is clear, looking at several colloquials (Mauritania, Morocco, Algiers, rural Tunisia (Takrouna), urban Tunisian (Bizerte), Egypt) that the system of conjunction in Arabic counts with a few elements that come from the classical language and have been assigned different tasks in the different languages:

a- the relative markers are all based on the article *il* with an increment -l + something, often a demonstrative.

b- the wh-words (characterized by their anaphoric function and their movement) are similar to the question words. Their main constituent is *aš*, the dubitative particle also found in negation.

c- there are the conjunctions always followed by a pronoun clitic with the verbs of saying *'an*/*inn*-, and the same *illi* that is widely recognized as a relative marker across the languages functioning as a simple conjunction in languages where the classical *ma* has been abandoned. This conjunction remains as a fossil form in Tunisian in:

b'ad ma	'after (that)'
ḳbal ma	'before (that)'

that I would compare to:

waḳt illi 'when, at the time that'

The whole system of conjunctions and prepositions would be an interesting topic of study in the modern colloquial forms of Arabic. It is a subject that fascinated the specialist, W. Marçais, who gave particular attention to it in his study of the Takrouna dialect.

16- Tunisian has backward pronominalization, as does English, which is restricted to the case of pronominalization down into a subordinate clause:

- 'ali y-'araf inn-u 'ali muriḍ
  'Ali knows that Ali is sick'

-*huwwa y-araf inn-u 'ali$_1$ muriḍ
  'he$_1$ knows that Ali is sick'

- waķt illi Dalila wasli.t li dar mta'-ha naḥḥa.t kabut-ha
  'when Dalila arrived to her house she took her coat' off

- waķt illi (hiyya)wasli.t li dar mta'-ha Dalila naḥḥa.t...
  'when she arrived to her house Dalila took off...'

Works Consulted

COHEN, D. 1963 — *Le dialecte arabe Hassaniya de Mauritanie*, Paris

HARREL, R. 1962 — *A short reference grammar of Moroccan Arabic*, Georgetown University, Washington, D.C.

LECOMTE, G. 1968 — *Grammaire de l'Arabe*, P.U.F., Paris

MARCAIS, W. and GUIGA, A. 1958 — *Textes arabes de Takrouna* (Tunisia), C.N.R.S, Paris

TAPIERO, N. — *Manuel de l'Arabe algerien*, Klincksieck, Paris

On Serbo-Croatian Complement Sentences

Colette Craig
March 16, 1972

This paper deals with verbal complement sentences in Serbo-Croatian.

It presents first the verbs which take an infinitival construction, regrouping them under headings of verbs of desire, modal-like verbs, verbs expressing the stages of a process and verbs of intention and realization.

Then the case of a group of verbs requiring a particular choice of aspect in their complement sentences will be introduced.

Considering next the question of the tenses of the embedded verbs, two kinds of embeddings will be defined, one which allows for all tenses and one which takes only the present tense.

Finally a note will be made of the way Serbo-Croatian handles the contrast between factive and non-factive complements by the choice of complementizer.

1.0 Let us first make some preliminary remarks on the status of the Serbo-Croatian language, which constitutes the source of data of this paper. The question of whether we are dealing with one or two "languages", or to be more specific, with two dialects of the same language or two separate languages, cannot and will not be resolved by linguistic expertise, for it has become a political issue that involves strong nationalistic feelings. But let us say that, for the purpose of this paper, an explicit distinction must be made between the two variants of Serbo-Croatian, the western-Croatian form and the eastern-Serbian form. The field work done for this study was conducted entirely in Zagreb, the capital of Croatia. The informants were all natives of Zagreb, and speakers of the standard štokavski jekavski dialect.

The area of syntax under observation, the verbal form of complement sentences, is one of the areas where the Serbian variant and the Croatian one happen to behave differently. The difference resides in the fact that Croatian has a more extensive use of the infinitival construction than Serbian does. So, while Croatian will prefer the construction:

(1) a. Želim ići
I want to go

Serbian will more normally use a complement sentence introduced by a complementizer da:

b. želim da idem
I want that I go
I want to go

Another example of this difference would be:

(2) a. počela sam zarađivati prije deset godina
begun I have to earn living before 10 years

(Croatian)

b. počela sam da zarađujem prije deset godina (Serbian)
I began earning my living ten years ago

Beside this dialectal variation along mostly geographical lines, one has to take into consideration the variation which exists between colloquial and standard (or literary) forms within Croatian:

(3) a. nastojim naći stan (Standard)

b. nastojim da nađem stan (Colloquial)
I am trying to find an apartment

What follows is a presentation of the verbs that command the infinitival construction in Croatian.

1.1 The first group of such verbs, which under no circumstances will take a da construction in Croatian if their subject is coreferential with the subject of the embedded verb, includes the verbs which express wishing or desiring. They are željeti: to want, voljeti: to like, and htjeti: to want[1]:

(4) želim govoriti s tobom
I want to talk with you

(5) volim {piti
{šetati po kiši

I like {to drink
{to walk in the rain

(6) on hoće zaboraviti svoju nesreću
He wants to forget his misfortune

The three verbs *nastojati*: *to strive*, *nadati se*: *to hope*, *źudjeti*: *to be eager*, which are related to this semantic group, were classified in this same group by most of the informants:

(7) žudim započeti s radom
I am eager to start with the work

(8) nastojim napisati to pismo
I am trying to write this letter

(9) nadala sam se završiti
hoping I was reflexive to finish
I was hoping to get finished

1.2 The second group is composed of modal-like verbs. They are best defined as corresponding to the French verbs *pouvoir*: *moći* (to be able to) and *devoir*: *morati* (to have to):

(10) možemo ići zajedno
We can go together

(11) moram otići zubaru
I have to go to the dentist

They can also be used in an impersonal form with *se*:

(12) može se reći...
can $\left\{\begin{matrix}\text{refl.}\\ \text{imp.}\end{matrix}\right\}$ say
One can say....

(13) mora se reći....
has $\left\{\begin{matrix}\text{refl.}\\ \text{imp.}\end{matrix}\right\}$ say
One has to say...

1.3 Partly comparable to these modal-like verbs are the verbs *trebati* and *valjati*, *to need* and *to be worthwhile*. There are two verbs *trebati* in fact: *trebati*$_1$ of:

(14) a. ja trebam knigu
I need a book

b. ja trebam raditi
I need to work

c. trebamo pronaći nove ljude
We need to find new people

d. trebamo doći u školu u osam sati.
We have to come to school at eight o'clock

This *trebati*$_1$ can appear in the impersonal construction with *se*:

e. treba se raditi mnogo brže ako se želi uspjeti
one needs to work much faster if one wants to succeed

and *trebati*$_2$ of:

(15) a. treba mi kniga
is necessary for me a book
I need a book

where *kniga* is the subject of treba and *mi* the indirect object

b. treba raditi
it is necessary to work

c. treba pohvaliti njegov rad
it is necessary to praise his work
His work should be praised

where *raditi* and *pohvaliti* are the subjects of *treba*.

To *trebati*$_2$ corresponds also the construction:

d. treba da radim
it is necessary that I work
I have to work.

where the *da* construction with its personal verb is the subject of *trebati*. In this construction a topicalization operation can apply that moves the subject pronoun of the embedded sentence up to the main clause:

(16) a. treba da *mi* pitamo iskusnije ljude

b. *mi* treba da pitamo iskusnije ljude
We should ask more experienced people

(17) a. treba da *vi* dođete u školu u osam sati

b. *vi* treba da dođete us školu u osam sati
You should come to school at eight o'clock.

Both *trebati*$_1$ and *trebati*$_2$ are heard in Zagreb. *Trebati*$_2$ is a more central, classical form of the verb, the one recommended by the grammarians of Croatian[2], and *trebati*$_1$ is a form more typical of the subdialect of Zagreb.

The verb *valjati*: *to be worthwhile* is also an impersonal verb which takes an infinitival complement:

(18) valja govoriti pošteno ako želimo doći
it is worthwhile talking honestly if we want to get

do cilja
to our goal
One should talk honestly if one wants to get to one's goal.

(19) ne valja biti prorok u svojoj zemlji
it is not worthwhile being a prophet in your own country
One should not be....
Il ne faut pas ....

Contrary to what happens to the impersonal *trebati*$_1$ in Croatian, *valjati* does not appear at all in a personal construction:

(20) a. valja da se brinemo o našoj omladini
it is needed that refl. we take care of our youth
(We) need to take care of our youth

b. *valjamo brinuti se.....

(21) a. valja da raščistimo ta pitanja
it is needed that we solve these questions
(We) need to solve these questions

bb. *valjamo raščišćavati.....

1.4 The third group is composed of verbs expressing the different stages of a process: početi: to begin, prestati: to stop, nastaviti: to continue.[3]

(22) počeo je zarađivati prije deset godina
begun has to earn his living before ten years
He began to earn his living ten years ago.

(23) jučer je prestao pušiti
yesterday has stopped to smoke
He stopped smoking yesterday

(24) nastavio je pjevati
continued has to sing
He continued to sing. He went on singing

1.5 The fourth group taking infinitival complement sentences could be partly described as expressing intention and accomplishment:[4]

(25) odlučio je čekati
He decided to wait

(26) namjerava otputovati
He intends to take off (leave)

(27) obećala sam kasno ustajati
promised I have late to get up
I promised to get up late

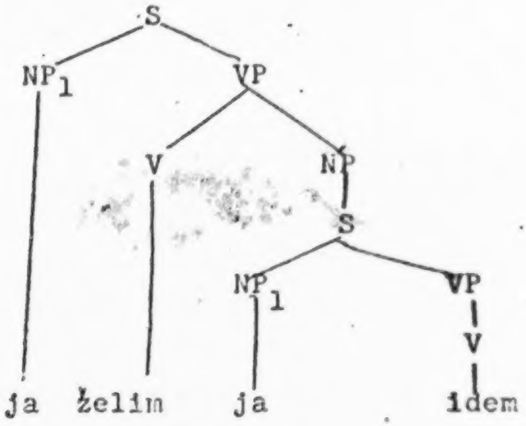
(28) uspjela sam uloviti vlak
I (F) succeeded in catching the train

(29) dospjeli su je vidjeti
managed they have her to see
They managed to see her.

(30) stigla sam naučiti lekciju
I (F) had enough time to learn my lesson (and did learn it)

(31) navikla sam kasno ustajati
I (F) am used to getting up late

1.6 In all the groups of verbs so far, the infinitive construction can be seen as the result of the application of an Equi-NP deletion rule applying only in case of coreferentiality of the subjects:



(32) a. ja želim ići
I want to go

If the subject of the main clause is not coreferential with the one of the embedded clause, a rule of complementizer insertion operates:

b. želim da (ti) ideš
I want that you go. I want you to go

c. *želim ti ići
Nom.

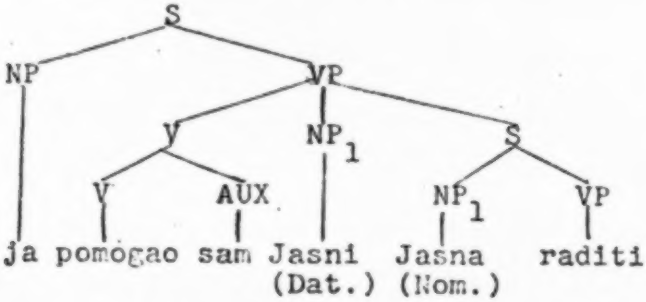
d. *želim te ići
Acc.

1.7 There are a few cases where the subject of the embedded clause, which is coreferential with the object of the main clause, undergoes an Equi-NP deletion rule. This happens with the two verbs pomagati: to help, and naučiti: to teach:

(33) pomogao sam Jasni raditi
helped I have Jasna (Dat.) to work
I helped Jasna to work

(34) naučio sam ga plivati
taught I have him (Acc.) to swim
I taught him to swim

Their corresponding structures are:



The same Equi-NP deletion rule applies with the verbs of perception, but the embedded clause shows up as a da construction then:

(35) čujem Ivana da pjeva
I hear Ivan (acc.) that he sings
I hear Ivan sing

(36) vidim Jasnu da pleše
I see Jasna (acc.) that she dances
I see Jasna dance

The infinitival construction exists also in a more literary form:

(37) čujem Ivana pjevati
I hear Ivan sing

(38) vidim Jasnu plesati
I see Jasna dance

2.0 Croatian, like other Slavic languages, has the characteristic of distinguishing between two verbal aspects, the imperfective and the perfective. With a few exceptions, all verbs fall into either one of the two aspect categories. So one identifies the verb *gledati*: *to look at* as an imperfective, and the verb *sjesti*: *to sit down*, as a perfective. The language also has a combination of derivational processes by which to turn an imperfective into a perfective and vice-versa.

2.1 Imperfectives are changed into perfectives by addition of a prefix, which generally narrows the meaning:

Imp:	piti:	to drink	*Perf*:	potpiti:	to drink up
	gledati:	to look at		pogledati:	to have a look
	plakati:	to cry		zaplakati:	to burst into tears

If one wants to study whether certain main verbs require a particular aspect in their complement sentences, and wants to isolate a command that could be written in terms of a (syntactic) feature like "+imperfective+, one has to disregard the primary imperfective verbs, the perfective correspondant of which undergo a change in meaning. A pair of perfective/secondary imperfective verbs like

pisati: imperfective$_1$ to write
potpisati: perfective to sign
potpisivati: imperfective$_2$ to sign

will permit us to consider a more straightforward syntactic fact.

2.2 One group of verbs was observed to impose a particular aspect on their verbal complement.[5] It is the group of verbs marking the different stages of a process:

(39) počela sam { prodavati (I2) } svu moju imovinu
{ *prodati (P) }
I began to sell all my belongings

(40) počeo je { kupovati (I2) } stare satove prije deset godina
{ *kupiti (P) }
He began to buy old clocks ten years ago

(41) prestao je { spominjati (I2) } njeno ime
{ *spomenuti (P) }
He stopped mentioning her name

(42) on mu je prestao { pomagati (I2) }
{ *pomoći (P) }
he him stopped to help
He stopped helping him

(43) nastavio je { potpisivati (I2) } indekse
{ *potpisati (P) }
He continued to sign indexes (students' reports)
He went on signing indexes

(44) nastavila sam { zavaravati (I2) } protivnika
{ *zavarati (P) }
I continued to mystify the opponent

This constraint on the aspect of the verbal complement is to be construed in terms of a semantic feature of the main verbs. It is not unlike what happens in other languages:

?*Il a commencé à naitre la semaine dernière.

?*He began to be born last week

Il a commencé à marcher la semaine dernière.

He began to walk last week.

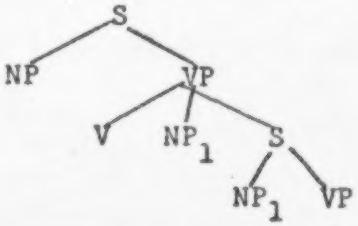
3.0 When considering the possible combinations of tenses that can be found in embedded clauses, a distinction has to be made between two kinds of constructions. One construction accepts all tenses; the other one is limited to the present tense.

3.1 Let us consider first the complement sentences in which all tenses can appear. All these verbs fall in the category of verbs of reported speech and thought, like kaziti: to say, dokazivati: to argue, očekivati: to expect, istićati: to point out, misliti: to think, znati: to know, shavaćati: to realize, etc. The way the reported part of speech is embedded in Croatian is by simple juxtaposition of what could be an independent clause. Therefore whatever is possible in an independent clause is also possible in an embedded clause. This holds for the reverse situation, too, which explains the fact that no present perfective verbs are to be found in embedded clause since they never appear in independent clauses:

(45) a. mislim da uvijek dolazi kasno — Imperfective present
b. je uvijek dolazio kasno — past
c. uvijek će dolaziti — future
I think that he always comes late
d. *dođe kasnije — Perfective present
e. *je dosao prekasno — past
f. *će doći prekasno — future

(46) a. kaže da pada kiša — Imperfective present
b. je padala kiša — past
c. će padati kiša — future
He says that the rain is falling
d. *napada kiša — Perfective present
ḋ. *je padala kiša u Zagrebu — past
f. *će napadati kiša u Zagreb — future

3.2 The da construction in which only the present is used corresponds to the following deep structure:



The subject of the embedded clause is coreferential with the object of the main verb.

(47) Kiša nas je sprečila da odemo P
rain us has kept that we go out
The rain kept us from going out

(48) podsjekio me je da ne zaboravim
reminded me has that not I forget
He reminded me not to forget

(49) on nam je predložio da idemo u kino I
he us has suggested that we go to the movies
He suggested (us) that we go to the movies

(50) spriječavala sam ga da ne puši I
kept I have him that not he smoke
I (fem.) kept him from smoking

(51) dozvolili su nam da to uradimo P
allowed they have us that that we do
They allowed us to do that

(52) {moliti ću / ask I will / molila sam / ask I have} ga da mi posudi kola
him that me he lends a car
I {will ask / have asked} him to lend me a car

(53) {navila / nagovorila} sam ga da ostane
persuaded I have him that he stay
I persuaded him to stay

(54) naređujem mu da dođe P
I order him to come

(55) on nam je preporučio da ne putujemo
he us has advised that not we leave
He advised us not to leave

(56) zatražio sam ga da ostane
begged I have him that he stay
I begged him to stay

The object of the main verb can also be found in a prepositional phrase.

(57) tražio je od oca da mu kupi odjelo
he demanded from his father that him buy suit
He asked his father to buy him a suit

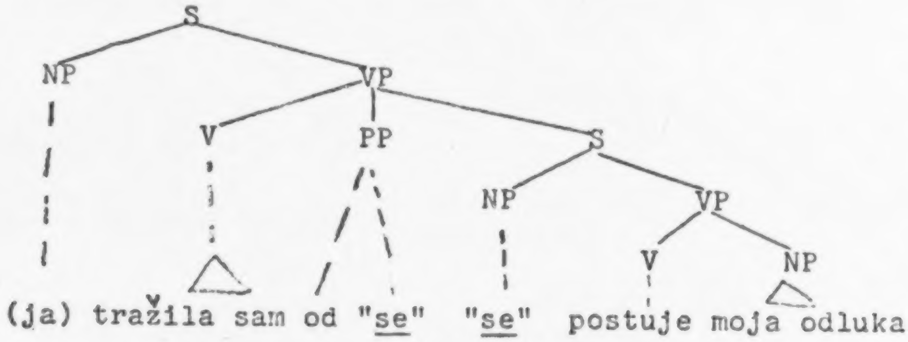
When the object is the impersonal se, it gets deleted from the upper sentence (taking with it the preposition).

(58) a. tražila sam da se poštuje moja odluka -Present
I demanded that "on" respect my decision
I demanded that my decision be respected

b. *tražila sam da se je poštivala moja odluka-Past

c. *tražila sam da će je poštivati moja odluka-Future

Corresponding to the tree:



3.3 Observe that both presents, the imperfective and the perfective one, can be found in these constructions. So, while the present perfective is ungrammatical in independent clauses like

*on dođe <doći^{Perf} <ići^{Imp}
come

*on ode <oći^{Perf} <ići^{Imp}
go out go

and in reported speech:

*kaže da dođeš
he says that you come

*kaže da odeš
he says that you go out

it is an acceptable form in the construct here:

(47) kiša nas je sprečila da odemo
The rain prevented us from going out

(54) naređujem mu da dođe
I order him to come

The case has been made that the present perfective of Slavic languages is ungrammatical and unobtainable on semantic grounds.[6] The perfective aspect of a verb cannot be expressed in the very open-ended present tense. The fact that it appears in these embedded clauses indicates that this present tense, always found in clauses embedded in a particular construction, is a cover-up tense in the same way the infinitive is a neutralized tense in English.[7]

4. The way Serbo-Croatian marks the difference between factive and non-factive complement sentences is through the choice of complementizer. As was presented in Bibovic's article (1971: B.3) the complementizer što corresponds to factive verbs, as in:

(59) žao mi je što sam poslušao njegov savet
(it is regretful to me...)
I regret that I followed his advice

(60) ne svidja mi se što je toliko lijen
(it does not please me...)
I dislike it that he is so lazy

(61) radovao se *što* je dobio prvu nagradu
He rejoiced *that* he had won the first prize

(62) vriječa se *što* je nitko ne shvaća ozbiljno
She resents *it* that nobody takes her seriously
feels offended

(63) sretan sam *što* te vidim
I am glad *to* see you

(64) krivo mu je *što* je Ivan bogat
(it is a *matter* of resentment to him...)
He resents John's being rich

The use of *što* is more widespread in Serbian than in Croatian but is still functional in Croatian.

Observe how the choice of complementizer in the following examples can make explicit whether the complement sentence is a factive or non-factive one:[8]

(65) a. ne volim *što* mi se mješaju u posao
I don't *like* their interfering in my business:
factive: they did

b. ne volim *da* mi se mješaju u posao
non-factive: if they did

(66) a. drago mi je *što* te vidim
I am glad *to* see you

b. Uvijek mi je drago *da* te vidim
I am always glad *(if)* to see you

Footnotes

1. Htjeti is also used as an auxiliary in the formation of the future. In this case it appears in its short form:

ću ići: I will go
ćeš govoriti: You will talk

hoću: I want
hoćes: You want

This short form is also the one the verb to want takes in a negative sentence. The negative form is therefore potentially ambiguous:

neću... I will not (future + neg.)
I do not want (present + neg.)

But in this case the verb to want will be followed by a da construction:

neću ići: I won't go

neću da idem: I do not want to go

Serbian, which has a future in ću + da has the ambiguous sentence:

$$\text{neću da idem: } \left\{ \begin{array}{l} \text{I will not go} \\ \text{I do not want to go} \end{array} \right.$$

2. Cf. Matica Hrvatska (institution functioning as the national Croatian academy) is explicit about the "ungrammaticality" of the infinitival construction:

ne treba da ideš onamo
You do not need to go that way

*ne trebaš ići onamo

Lj. Jonke (1965: 396) also notes as incorrect the form:

*trebate doći

and gives as the only acceptable one:

vi treba da dođete

3. But the verb svršiti: to finish does not fall into this category. It takes a verbal noun instead of an infinitive.

svršio je s pisanjem knjige
He finished (with the) writing (of) the book

oni će završiti s ispitivanjem đjaka kroz
They will finish examining the pupils in
(with the examination of)
pola sata
half an hour

4. This verb obećati: to promise and the following nadati se: to hope show an alternation between the infinitive and the da construction. In the past, they take the infinitive as seen in the example, but in the present they tend to take a verb in an overt future form.

?obećajem doći

obećajem da ću doći u večer
I promise that I will go tonight

as opposed to

obećala sam doći
I promised to go

and

?nadam se završiti u večer

nadam se da ću završiti u večer
I hope that I will finish tonight

nadala sam se zavrišiti
I was hoping to finish

5. Another interesting distribution of aspects was observed in the negative imperative sentences. Serbo-Croatian, unlike Russian, has two negative particles, ne and nemoj. Russian only has ne, which is always followed by the imperfective as-

pect, which is also the case in Croatian. Nemoj on the other hand, can take either aspect.

nemoj otvoriti bocu Perf.

nemoj otvarati prozor Imp.
Don't open the {bottle
window

ne otvaraj prozor Imp.

*ne otvori {prozor Perf.
bocu

Serbian has:

nemoj da dođete Perf.
Don't come

nemoj doći prije osam Perf.
Don't come before 8:00

ne dolazi prije osam Imp.

nemoj to sada početi Perf.
Don't start this now

ne počinji opet Imp.
Don't start over again

ne prelazi cestu Imp.
Don't cross the road

*ne prijeći cestu Perf.

nemoj prelaziti cestu Imp.

nemoj prijeći cestu Perf.
Don't cross the road

6. Actually, the present perfective is sometimes used in Croatian with a subjunctive or conditional value.

7. Note the parallel between the present perfective and the infinitive in Croatian itself in the sentences:

mislim da je trebao { da se požurimo Perf.
                     { žuriti se Imp.
I think that it is necessary that we hurry

mislim da je potrebno { da se požurimo Perf.
                       { žuriti se Imp.

which the informants preferred to the complementary forms

mislim da je trebalo požuriti se
            je potrebno da se žurimo

8. Wayles Browne pointed out to me that the factive *što* corresponds to the *da* + all tenses construction, while the non factive *da* introduces a construction with the neutralization of tenses in the present.

Works Consulted

Brabec, Hraste, Živković

Gramatika Hrvatskosrpskoga jezika

Zagreb 1955

Bibovic Ljiljiana

Some remarks on the factive and non-factive complements in English and Croatian

Institut za lingvistiku B studies 3
Zagreb 1971

Jonke ljudevit

Književni jezik u teoriyi i praksi

Zagreb 1965

Matica Hrvatska

Jezični savjetnik s gramatikom

Zagreb 1971

Meillet A.
Vaillant A.

Grammaire de la langue Serbo Croate

Paris 1952

Outline of Bengali Grammar

Sheila Jasanoff

Transcription: The symbols used are essentially self-explanatory. The dotted series (ṭ, ṭh, etc.) is alveolar rather than retroflex, ɛ, ɔ are low mid vowels, and bh, dh, etc., are voiced aspirates. /s/ is pronounced [ʃ] except before dentals.

I. The Noun

1. Human nouns and pronouns are marked for both number and case. The four cases are the direct, oblique, genitive and locative.

	Direct	Oblique	Genitive	Locative	
Pronoun					
Sg.	o	oke	or	ote	o='he'
Pl.	ora	oder	oder	--	
Noun					
Sg.	chele	cheleke	cheler	chelete	chele='boy'
Pl.	chelera	cheleder	cheleder	--	

2. Non-human pronouns may be either abstract or concrete and are declined as follows:

	Direct	Oblique	Genitive	Locative
Abstract	ta	take	tar	tate
Concrete				
Sg.	(eṭa, oṭa, seṭa)	(e-, o-, seṭake)	(e-, o-, seṭar)	(e-, o-, seṭate)
Pl.	egulo, ogulo, segulo	e-, o-, seguloke	e-, o-, segulor	e-, o-, segulote

The concrete pronouns consist of the demonstrative adjectives *e* 'this', *o* 'that' and *se* 'that (remote)' plus the definite article *ta* in the singular and in the plural the collective article *gulo*.

3. Non-human nouns are only optionally marked for number (using the collective article *gulo*, e.g., *boi* 'book(s)', *boigulo* 'books'). Non-human nouns are not generally marked as obliques but are so marked optionally when they are definite and obligatorily when they appear in constructions in which they have a complement. Thus

boiṭa(ke)	rakho	'Keep the book'
book the	keep	

eṭake	ki	bɔle	'what does one call this'
this	what	calls	

ami	ei	pathortake	hire	mone kori	'I consider this stone a
I	this	stone	diamond	consider	diamond'

II. The Verb

1. Primary Verbs

1.1. Finite Forms: The verb in Bengali is inflected for person but not number. There are two verb classes, the simple stem class (Class I) and the augmented stem class (Class II). Only third person forms are given in the paradigms below:

	Class I: Stem Kɔr--'do'		Class II: Stem chɔra--'scatter'	
Present	o kɔre	'he does'	o chɔrae	'he scatters'
Present Imperfect	o korche	'he is doing'	o chɔraeche	'he is scattering'
Present Perfect	o koreche	'he has done'	o chorieche	'he has scattered'
Past	o korlo	'he did'	o chɔralo	'he scattered'
Past Imperfect	o korchilo	'he was doing'	o chɔracchilo	'he was scattering'
Past Perfect	o korechilo	'he had done'	o choriechilo	'he had scattered'
Past Frequentative	o korto	'he used to do'	o chɔrato	'he used to scatter'
Future	o korbe	'he will do'	o chɔrabe	'he will scatter'
Present Imperative	ko uk	'let him do'	chɔrak	'let him scatter'

1.2. Non-Finite Forms

Verbal Noun -- This is the citation form of the verb.

kɔra 'doing' chɔrano 'scattering'

The alternate forms *korba-*, *chɔraba-* do not regularly occur without a case marker, e.g., *korbar* 'of doing'.

Past Participle

Passive	*kɔra*	'done'	*chɔrano*	'scattered'

This participle coincides formally with the verbal noun.

Active	*kore*	'having done'	*chorie*	'having scattered'
Infinitive	*korte*	'to do'	*chɔrate*	'to scatter'
Present Participle	*korte*	'doing'	*chɔrate*	'to scatter'
Conditional Participle	*korte*	'if (one)does'	chɔrate	'if (one) scatters'

This participle qualifies a noun or pronoun of any person in the direct case: *o korte* 'if he does', *ami korte* 'if I do'. Uses of these forms will be discussed in the section on syntax.

2. Causatives

Causative verbs constitute the only synthetic derived category in Bengali. They are formed productively only from verbs of Class I. The causative stem is derived by adding -*a* to the stem of the primary verb (-*wa* when the latter ends in a vowel), e.g., *kɔr*- 'do', *kɔra* 'cause to do', *kha*- 'eat', *khawa*- 'cause to eat'. Causatives are conjugated like verbs of Class II.

3. Compound Verbs

These are of three principal types.

3.1. Noun and Primary Verb

Many very common verbs in Bengali consist of one of a highly restricted set of primary verbs (particularly kɔra 'to do', dewa 'to give', khawa 'to eat') plus a direct object. Thus

bɛbohar 'use (n.)' + kɔra = bɛbohar kɔra 'to use'
bhɔe "fear (n.)' + khawa = bhɔe khawa 'to get frightened'

Verbs of this sort tend to function like single lexical items. For instance, a transitive compound verb like bɛbohar kɔra can take a direct object like any transitive primary verb, e.g., boi bɛbohar kɔra 'to use a book'. Apart from causative formation (discussed above), compounding nouns with kɔra is the closest thing Bengali has to a productive process of verb derivation. An interesting extension of this process is the borrowing of verbs from English. The unmarked English verb fills the slot of the noun, and examples like phon kɔra 'to telephone' and maind kɔra 'to mind' are extremely common in colloquial Bengali.

3.2. Past Active Participle and Primary Verb

In the preceding class of compound verbs the primary verbs kɔra, khawa, etc., are essentially meaningless elements serving only to derive verbs from existing nouns. In this class, on the other hand, the primary verb functions as an aspect marker (generally perfective) in more or less the same way as English prepositions in verbs like 'to eat up', 'to run off'. Thus:

Khawa 'to eat' and phɛla 'to drop' ⇒ khee phɛla 'to eat up'

harano 'to lose' and jawa 'to go' ⇒ harie jawa 'to get lost'

These formations are in a sense idiomatic, since the meaning of the compound verb is not predictable from the lexical meaning of its components, but there is a certain amount of regularity in that all verbs with phɛla, for instance, have the same connotation of completed action. This type of composition is a special case of verb concatenation in conjoined sentences (see below).

3.3. Noun and Primary Verb (Impersonal)

Many verbs of feeling, sensation, etc., belong to this class. They are used in impersonal constructions in which what would be the grammatical subject in English appears in the genitive case. Thus we have the verb mathabɛtha kɔra 'to have a headache' (=matha 'head' and betha 'pain' and kɔra 'do') in a sentence like:

amar	mathabɛtha	kɔrche	'I have a headache'
of me	headache	does	

The syntactic relations that obtain among the constituents in sentences of this sort is not transparent and will be reserved for later discussion. In particular, it must be noted that the constituent structure of the sentence above is represented by the labelled bracketing:

$$_{s}[\,_{np}[\text{amar}]_{np}\ _{vp}[\,_{np}[\text{mathabɛtha}]_{np}\,_{v}[\text{korche}]_{v}\,]_{vp}\,]_{S}$$

amar is therefore not a possessive genitive modifying *mathab tha*.

III. *The Noun Phrase*

The possessive genitive comes initially in the NP and the case marker comes finally. The underlying representations of these elements will not be treated here. The order of the remaining items in the NP depends on whether the NP is definite or not. What follows can be formulated more generally in generative terms but will be presented for the time being within the structural framework of this sketch: The definite NP expands as follows:

$$\left\{\begin{array}{l}\text{(Demonstrative)}\\ \\ \text{(Relative)}\end{array}\right\}\ \text{(Adjective)}\ \ \text{Noun}\ \ \text{(Numeral)}\ \left\{\begin{array}{l}\text{Article (ṭa, \underline{gulo}, etc.)}\\ \\ \text{Classifier (\underline{jon}, \underline{khana}, etc.)}\end{array}\right\}$$

ei	bɔro	boi	car	ṭa
this	big	book	four	"the"

'these four big books'

In the indefinite NP the order is as follows:

(Numeral	Classifier)	(Adjective)	Noun	
car	ṭa	bɔro	boi	'four big books'
four	"the"	big	book	

IV. The Verb Phrase

Adverbs precede the verb, normally in the order time, manner, place. The order of adverbs is not fixed with respect to objects.

1. Manner Adverbs

These can be either single morphemes (*aste* 'slowly') or an adjective and *kore* 'having done' (*bhalo* 'good' and *kore* = *bhalo kore* 'well').

2. Place Adverbs

Many of these are locatives of nouns denoting spatial relationships, thus *opore* 'above' from *opor* 'top'. These locatives function as post-positions governing the genitive case and are used to form longer adverbial phrases, e.g., *chader opore* 'on the roof'.

3. Time Adverbs

These again can be single morphemes (*ɛkhon* 'now') or post-positional phrases (*chɔṭar pɔre* 'after six').

V. The Sentence

1. The Simple Sentence

1.1. Declarative Sentence

The normal word order in the Bengali declarative sentence is:

Subject	Indirect Object	Direct Object	Verb
ami	Ranake	boi	debo
I	to Rana	book	will give

'I will give Rana a book'

1.2. Nominal Sentence

In the present tense the nominal sentence contains no copula:

Rana	bhalo	chele	'Rana is a good boy'
Rana	good	boy	

In the past tense the copula is formed from the stem chi- (3rd person chilo 'was') and in the future from the stem hɔ- (3rd person hɔbe 'will be').

1.3. Imperative Sentence

There are special imperative forms for the second and third persons. In the second person there are different forms for the present and the future imperative. The subject of an imperative is generally deleted:

dekhi	'let me (us) see'
kɔro	'do (now)'
koro	'do (in the future)'

1.4. Impersonal Sentence

1.4.1. One kind of impersonal sentence has been discussed above in connection with the third type of compound verb. Another illustration is given below:

amar	kanna	peeche	'I feel like crying'
of me	crying	has found	

1.4.2. A second type of compound sentence involves a logical subject carrying out an action out of necessity. The necessary action is expressed by means of the infinitive and third person forms of hɔwa 'to be, become'. The logical subject appears either in the genitive or in the oblique:

$$\left\{\begin{array}{l}\text{amar}\\ \text{of me}\\ \\ \text{amake}\\ \text{me}\end{array}\right\}\quad \begin{array}{ll}\text{jete} & \text{hɔbe}\\ \text{to go} & \text{will be}\end{array}\qquad \text{'I will have to go'}$$

1.5. Passive Sentence

Bengali has no regular way of forming a passive from an active sentence. There is an impersonal passive construction using the past passive participle and the verb hɔwa 'to be' or jawa 'to go'. The agent can be expressed as a genitive when the verb is *hɔwa* and is in a perfect tense. No agent can be expressed when the verb is *jawa*.

boiṭa	pɔra	hɔe	'the book gets read'
the book	read	becomes	

boiṭa	amar	pɔra	hoeche	'the book has been read by me'
the book	of me	read	has become	

boiṭa	pɔra	jae	'the book can be read'
the book	read	goes	

In a similar construction using the past passive participle of a causative verb the agent is expressed in a postpositional phrase using the postposition *die* 'by, through':

Boiṭa oke die pɔrano jabe	'The book can be made to be read by him'
Boiṭa oke die pɔrano hɔbe	'The book will be made to be read by him'

1.6. Negative Sentence

The negative particle *na* follows the finite verb to form the negative of all tenses except the present perfect and the past perfect

ami kori, korbo, kortam na	'I don't, won't, didn't use to do'

To form the negative of the perfect tenses the particle *ni* is added to the present tense of the verb:

tumi korecho and na ⇒ tumi kɔro ni	'you haven't done'
tumi korechile and na ⇒ tumi kɔro ni	'you hadn't done'

1.7. Question

1.7.1. Yes-No Question

Question words begin with *k-*, thus *ki*, *ke*, *keno*, *kothae* 'what, who, when, where'.

A sentence may be questioned in two ways, either by marking the simple declarative sentence with question intonation (rising on final syllable) or by using the unstressed interrogative particle *ki* either before or after the verb as shown:

tumi korecho?	
tumi ki korecho?	'have you done?'
tumi korecho ki?	

1.7.2. Wh-Question

To question a specific constituent of a sentence one replaces the item to be questioned with the corresponding interrogative. The structure of the sentence remains unchanged:

tumi ki khao?	'what do you eat?'
tumi kothae thako?	'where do you live?'

Note that the interrogative pronoun *ki* is always stressed.

2. The Complex Sentence

2.1. Conjoined Sentences

Sentences may be conjoined using conjunctions *ar* 'and', *kintu* 'but', *karon* 'because', etc.

ami asbo ar tumi asbe	'I'll come and you'll come'
ami asbo kintu tumi asbe na	'I'll come but you won't'
ami asbo karon tumi asbe	'I'll come because you will'

Where the clauses to be conjoined have identical subjects and the conjunction is *ar*, the following construction using the past participle active and denoting sequential action is more usual:

ami	*bari*	*phire*	*juto*	*khule*	*khater*
I	home	having returned	shoe	having taken off	of the bed

upor	*bose*	*boi*	*pori*
on top	having sat	book	read

'I get home, take off my shoes, sit on top of the bed and read a book.'

Note that in such a sentence the tense of the sequence is determined entirely by the finite verb. This construction is, of course, analogous to the one described below in which the conditional participle denotes a temporal sequence rather than a condition. The difference is that the past participle type is reserved for clauses with the same subject. (Compound verbs of type 3 are a special case of this kind of conjunction.) A clause expressing cause may precede the clause expressing effect, but in that case the petrified participle *bole* 'because' (literally 'having said') must be used in place of *karon* and after the verb.

ami asbo bole ora aste caeni
I will come because they to come did not want
'They didn't want to come because I'm coming.'

nek kaj koreche bole ora klanto
much work have done because they tired
'They're tired because they've done a lot of work.'

*karon ami asbo ora aste caeni

2.2. Relative Sentences

2.2.1. Non-restrictive Relative Clauses

The noun to be relativized has placed before it the relative pronoun *je* which, as stated above, patterns like a demonstrative. The coreferential noun in the main clause is pronominalized to the syntactically appropriate form of the pronoun *se*. The relative clause generally precedes the main clause.

je	lokti	eseche	se	ekhon	khacche
what	man	has come	he	now	

'The man who came is now eating.'

2.2.2. Restrictive Relative Clauses

The main clause remains unchanged. The noun to be relativized is replaced by the pronoun je 'who' and the clause is inserted after the antecedent noun in the main clause.

lokti,	jake	ami	kal	dekhechi,	khacche
the man	whom	I	yesterday	saw	is eating

'The man, whom I saw yesterday, is eating.'

2.2.3. Relative Adjectives and Adverbs

These occur in pairs and behave in much the same way as relative pronouns:

tumi	jɔkhon	jabe	amio	tɔkhon	jabo
you	when	will go	I too	then	will go

'When you go I'll go too.'

tumi	jɔto	phɔl	khabe	amio	tɔto	phɔl	khabo
you	as much	fruit	will eat	I too	that much	fruit	will eat

'I'll eat as much fruit as you.'

Bengali also has the indefinite relative pronouns *je keu* 'whoever', *ja kichu* 'whatever'.

2.3. Indirect Discourse

2.3.1. Reported Declarative Sentences

In reported speech the tenses remain the same as in direct speech. The pronouns in the reported clause are altered when necessary. The reported clause is optionally introduced by the conjunction *je* 'that':

o	boleche	(je)	kal	asbe
he	has said	(that)	tomorrow	will come

'He said he'd come tomorrow.'

2.3.1. Reported Questions

To form an indirect question the interrogative particle *ki* is replaced by the compound particle *kina* 'whether' which follows the verb:

o	jiges	korlo	ami	asbo	kina
he	question	made	I	will come	whether

'He asked whether I'd come.'

2.4. Conditional Sentences

Conditional sentences may be formed by using the conjunction *jodi* 'if' and the resumptive *tɔbe*, *ta hole* 'then, in that case', or by using the conditional participle:

o	jodi	ase	tɔbe	ami	cole	jabo
he	if	comes	then	I	will	go away

o	ele	ami	cole	jabo
he	if comes	I	will	go away

'If he comes I'll go away.'

(The second sentence above is ambiguous in that it can also mean 'After he comes I'll go away.') The word order shown in the first sentence is the usual one, but the position of *jodi* is not fixed. In negative conditional clauses the particle *na* precedes the verb:

o	jodi	na	ase	ta hole	ami	asbo	na
he	if	not	comes	then	I	will	not come

'I won't come if he doesn't come.'

In either type of conditional clause forms of the root *ach-* 'be' are replaced by forms of *thak-* 'stay':

o ache 'he is (there)' but o jodi thake 'if he is (there)'

This extends to the perfect tenses in which the root *ach-* (*ch-*) appears as part of the finite form, so that:

o eseche 'he has come' but o jodi ese thake 'if he has come' not *o jodi eseche.

There is no corresponding paraphrase with thak- for the continuous tenses and these simply may not appear in a conditional sentence.

2.5. Temporal Clauses and Assorted Embeddings

Embedded temporal clauses belong to the general class of embeddings in which the finite verb of the subordinate clause is replaced by a non-finite form. In the case of temporal clauses the non-finite form in question is the verbal noun, whose case is determined by the particular temporal postposition used. The embedded clause can appear initially, immediately after the subject or finally. The subject of an embedded temporal clause is obligatorily deleted:

kolkata	jaoar	pɔr	Rana	amake	cithi	likhlo
Calcutta	of going	after	Rana	me	letter	wrote

'After going to Calcutta Rana wrote me a letter.'

Also: Rana kolkata jaoar pɔr amake cithi likhlo

Rana amake cithi likhlo kolkata jaoar pɔr.

*o kolkata jaoar pɔr Rana amake cithi likhlo etc.

jaoar age 'before going', jaoa obdhi 'until going', jaoa maltro 'as soon as going'. Other embeddings using non-finite forms are illustrated below:

(1) Verbal Noun -- Nominative or Direct

tomar	erokom	kɔra ṭa	amar	bhalo	lageni	
your	this way	doing	of me	good	did not strike	

'Your acting this way didn't please me.'

Oblique -- here the verbal noun is a complement

tomar	erokom	kɔrake	ami	bhalo	mone kori	na
	this way	doing	I	good	consider	not

'I don't consider your acting this way good.'

Locative

tumi	erokom	kɔrate	ami	khusi	hoini
you	this way	in doing	I	pleased	was not

'I wasn't pleased at your acting in this way.'

Genitive

tomar	okhane	jaoar	dɔrkar	nei
of you	there	of going	need	is not

'There is no need for you to go there.'

ami	kaj	korbar	jonno	ekhane	esechi
I	work	of doing	for purpose	here	have come

'I've come here to work.'

(2) Infinitive

ami ekhane kaj korte esechi
I here work to do have come
'I've come here to work.'

Reflexives:

The most common way of forming reflexives is by using the reflexive pronoun nij 'self' which is unmarked for person.

nijeder kaj ses kore ora bari cole gelo
own work having finished they home went away
'Having finished their own work they went home.'

o nijeke dekhlo
he himself saw
'He saw himself.'

Number in the Bengali NP

Sheila Jasanoff

The Bengali finite verb is not marked for number. Below, I shall discuss the ways in which number can be indicated in the NP.

1. As has been stated elsewhere, human nouns and pronouns have both a singular and a plural inflection. Thus:

	Singular		Plural	
Direct	chele	'boy'	chelera	'boys'
Oblique	cheleke		cheleder	
Genitive	cheler		cheleder	
Locative	chelete			

The plural noun can be used either indefinitely (generically) or definitely.

chelera	mach	bhalobase	
boys	fish	like	'boys like fish'

chelera	kothae	geche	
boys	where	have gone	'where have the boys gone'

The status of the plural is complicated by usages like the following:

(i) Ranara ('the Ranas'), 'Rana and others'
(ii) mare 'mothers', 'mother and others'
(iii) mababara 'mothers and fathers', 'mother, father and others' or 'mother and father'

In Bengali, as in English and most western languages, the plural can indicate a collection of like objects, e.g., mara can mean 'several mothers'. More interestingly, however, a plural ending on a proper name or kinship term can indicate that the person in question is part of a larger set of unlike members whose nature is to be deduced from the context. Thus:

Ranara	pac	bhaibon	'Rana and his siblings are five brothers and sisters.'
Rana (pl.)	five	brother sister	

Ranara	kothae thake?	'Where do Rana and his family live?'
Rana (pl.)	where stays	

In English this kind of plurality is found only in the pronominal system where we means 'I and others' and you (pl.) means 'you and others' as well as 'you and you'. The plural of a proper name specifically excludes this kind of plurality, e.g., Henrys means 'Henry and Henry' in 'two Henrys ruled England from 1485-1547.'

It should be noted that example (iii) is triply ambiguous because a compound like mababa 'mother and father' can be redundantly marked as a plural.

2. An unmarked non-human noun can be interpreted as either singular or plural:

ekhane	kar	boi	pore	ache?	'whose book(s) is (are) lying here?'
here	whose	book	fallen	is	

When the noun is definite, number is indicated by using the definite article ta in the singular and the 'collective article' gulo in the plural. The following pair of sentences illustrates the use of gulo:

o	boi	kinbe	
he	book	will buy	'he will buy books'
o	boigulo	kinbe	
he	the books	will buy	'he will buy the books'

Unlike the nouns, non-human pronouns obligatorily contain reference to number as shown below:

	Singular	Plural
Direct	eṭa, oṭa, seṭa	egulo, ogulo, segulo
Oblique	e-, o-, seṭake	e-, o-, seguloke
Genitive	e-, o-, seṭar	e-, o-, segulor
Locative	e-, o-, seṭate	e-, o-, segulote

It will be seen that these pronouns simply consist of the demonstrative adjective *e* 'this', *o* 'that' and *se* 'that (remote)' plus the definite article *ṭa* in the singular and the collective article *gulo* in the plural. We therefore get sentences like the following in which a plural pronoun resumes an unmarked noun:

o	boi	kine	seguloke	bari	nie	jabe
he	book	having brought	them	home	having taken	will go

'he will buy books and take them home'

3. A syntactic way of indicating number in the NP is supplied by adjective reduplication. In particular, adjective reduplication allows us to distinguish singular from plural for indefinite, non-human nouns:

o	sundor	boi	kinbe
he	beautiful	book	will buy

'he will buy {a beautiful book' / beautiful books'}

o	sundor	sundor	boi	kinbe
he	beautiful	beautiful	books	will buy

'he will buy beautiful books'

4. There are certain environments in which number is neutralized for all nouns. The commonest of these are listed below.

(i) After certain quantitative adjectives like *ɔnek* 'many' and *besi* 'more': *ɔnek chele* 'many boys', *ɔnek boi* 'many books'

(ii) After numerals: *duti chele* 'two boys', *duto boi* 'two books'

(iii) In apposition with pronouns:

ora amar bhai
they my brother 'they are my brothers'

ogulo amar boi
those my book 'those are my books'

Adjective reduplication can be used in these environments: *ɔnek sundor sundor chele*, 'many beautiful boys' and *duto sundor sundor boi*, 'two beautiful books'.

5. It would appear from the discussion above that although number can be indicated in the Bengali NP, it is not a simple grammatical category as in English. Neither verbs nor adjectives are inflected for number and only human nouns have an obligatory singular-plural distinction. Even for these, as has been seen in #1 above, what I have called "plurality" can be something other than a simple mark of number.

Pronominalization and Equi-NP Deletion
in Bengali

Sheila Jasanoff

1. Conjoined Sentences

1.1. If two sentences S_1 and S_2 are conjoined in that order and a noun N_1 in S_1 is coreferential with a noun N_2 in S_2, then N_2 may be deleted when N_1 and N_2 are in the same case.

1.1.1. N_1 and N_2 in Direct Case

(i) Rana hase ar Rana Khele Rana hase ar khele.
Rana laughs and Rana plays Rana laughs and plays.

In (i) N_2 may not be replaced by a pronoun; in other words (i') is ungrammatical when o refers to Rana:

(i') *Rana hase ar o khele.
Rana laughs and he plays.

1.1.2. N_1 and N_2 not in Direct Case

When N_1 and N_2 are not in the direct case, the deletion of N_2 is customary only if S_1 and S_2 are short:

(ii) ami Ranake daki ar Ranake khawai ami Ranake daki ar
I Rana call and Rana feed I Rana call and
'I call Rana (over) and feed Rana 'I call and
khawai. feed Rana.'
feed.

In (ii) N_2 may be replaced by a pronoun to yield (ii'):

(ii') ami Ranake daki ar oke khawai.
I Rana call and him feed.
'I call Rana and feed him.'

2. Complex Sentences

2.1. When S_2 is an embedded sentence in a matrix sentence S_1 and a noun N_1 in S_1 is coreferential with a noun N_2 in S_2, N_2 is obligatorily deleted if N_1 and N_2 are either in the direct case or in the genitive in certain types of constructions. (These genitive constructions will be described in detail elsewhere. It seems likely that the noun in the genitive case is an underlying subject.)

(i) *Rana Rana asar pɔre khabe ⇒ Rana asar
*Rana Rana of coming after will eat ⇒ Rana of coming

pɔre khabe
after will eat.
*'Rana after Rana coming will eat' ⇒ 'Rana will eat after coming.'

(ii) *Rana Ranar sorir kharap hoeche bole aste
Rana of Rana body bad has become since to come

parbe na ⇒
will be able not ⇒

'Rana will not be able to come since Rana is ill.' ⇒

Rana sorir kharap hoeche bole aste parbe na

Rana body bad has become since to come will be able not

'Rana will not be able to come, since (he) is ill'

2.2. In all other situations forwards pronominalization applies (optionally) in a complex sentence.

(iii) Rana asar p re ami Ranake khawabo ⇒ Rana asar pɔre ami oke khawabo.

Rana of coming after I Rana will feed ⇒

'After Rana comes I will feed Rana' ⇒ 'After Rana comes I will feed him.'

2.3. Backwards pronominalization is a highly restricted process in Bengali and takes place only when the following conditions are met:

a. S_2 (the embedded sentence) precedes S_1 (the matrix sentence).

b. N_2 is not in the direct case or the genitive case (see above).

c. N_1 is in the direct case or the genitive case (see above).

For example:

(iv) ami Ranake khawabar pɔre Rana bari gelo

I Rana of feeding after Rana home went

'After I fed Rana Rana went home.'

⇒ ami oke khawabar pɔre Rana bari gelo
I him of feeding after Rana home went
'After I fed him Rana went home.'

LANGUAGE-PARTICULAR RULES AND EXPLANATION IN SYNTAX*

David M. Perlmutter and Janez Orešnik

Part One -- An Analysis of the Orphan Accusative in Slovenian

1. The Accusative Prediction Rule

In Slovenian, masculine and feminine nouns and adjectives have the following forms in the nominative, accusative, and genitive singular.[1]

	Masculine Inanimate	Masculine Animate	Feminine
NOMINATIVE	navaden ječmen	navaden človek	navadna ajda
ACCUSATIVE	navaden ječmen	navadnega človeka	navadno ajdo
GENITIVE	navadnega ječmena	navadnega človeka	navadne ajde
	'ordinary barley'	'ordinary man'	'ordinary buck-wheat'

For feminines, the three cases are distinct. For masculines, however, the form of the accusative can be predicted by the following rule:[2]

(1) Accusative Prediction Rule

a. For animates, the accusative is like the genitive.

b. For inanimates, the accusative is like the nominative.

2. The Orphan Accusative

In answer to the question

(2) Katero ajdo hočete?
which buckwheat you want
'Which buckwheat do you want?'

one can answer

(3) Hočem navadno ajdo.
I want ordinary buckwheat
'I want ordinary buckwheat.'

or simply

(4) Hočem navadno.
'I want ordinary.'

with the head noun ajdo absent. Similarly, one could answer (2) with either

(5) Navadno ajdo.
'Ordinary buckwheat.'

or simply

(6) Navadno.
'Ordinary.'

with deletion of ajdo.

In answer to the question

(7) Kateri ječmen hočete?
'Which barley do you want?'

one can answer

(8) Hočem navaden ječmen.
'I want ordinary barley.'

parallel to (3). However, the counterpart of (4),

(9) *Hočem navaden.

'I want ordinary.'

is ungrammatical. Instead, one must say

(10) Hočem navadnega.

'I want ordinary.'

Similarly, while one can answer (7) with

(11) Navaden ječmen.

'Ordinary barley.'

one cannot say

(12)*Navaden.

'Ordinary.'

parallel to (6), but must answer instead

(13) Navadnega.

'Ordinary.'

When the head noun is not present, the adjective has the genitive form instead of the nominative form that is the usual accusative form for inanimates. Because the genitive-like accusative form navadnega appears in the absence of the head noun, we will refer to it as the Orphan Accusative.

The Orphan Accusative occurs not only with mass nouns like ječmen, 'barley', but also with count nouns. Thus, in answer to the question

(14) Kakšen površnik hočete?

what kind of overcoat you want

'What kind of overcoat do you want?'

one can answer

(15) Hočem navaden površnik.

'I want an ordinary overcoat.'

but not

(16)*Hočem navaden.

Instead, one must say

(17) Hočem navadnega.

'I want an ordinary one.'

Similarly, one could answer (14) with

(18) Navaden površnik.

'An ordinary overcoat.'

but not with

(19)*Navaden.

Instead, one would have to say

(20) Navadnega.

'An ordinary one.'

The key fact is that inanimate masculines have an Orphan Accusative form distinct from their ordinary accusative, while feminines do not.[3] The Orphan Accusative therefore cannot be treated as a partitive genitive or any other recognized grammatical category, since such a treatment would not explain why it occurs with masculines but not with feminines. Furthermore, while the other Slavic languages have paradigms like that in §1, in which the accusative of the masculine is like the genitive if it is animate and like the nominative if it is inanimate, the Orphan Accusative, as far as we know, is unique to Slovenian. The Orphan Accusative is a language-particular phenomenon, and we are faced with the question of how it is to be dealt with in an explicit grammar of Slovenian.

Using a rewrite rule of the type one finds in the linguistic literature, we could state the facts as follows:

(21) *Orphan Accusative Rule*

$$\begin{bmatrix}\text{+Adjective}\\ \text{+Masculine}\\ \text{+Accusative}\end{bmatrix} \longrightarrow [\text{+Genitive}]/\text{in the absence of the head noun}$$

We might then claim that we had accounted for the Orphan Accusative in Slovenian.

What is striking about the Orphan Accusative Rule is the fact that it states outright the class that it applies to (masculine accusative adjectives), the change that they undergo (they become genitive), and the environment in which this happens (in the absence of the head noun). Because the Orphan Accusative Rule *states* these facts, it does not *explain* them. It is the purpose of this paper to show that pseudo-rules like (21) serve only to make us aware of the facts that need to be explained. What is needed is not a rule that merely states the facts, but a solution that explains them.

3. Some Questions Raised by the Orphan Accusative

Examination of the Orphan Accusative Rule (21) immediately raises a number of questions that demand answers.

Question One. *Why is it that the special form found in the Orphan Accusative is not an arbitrary ending such as -uruburu or -gnuf, but rather an ending that exists elsewhere in the adjectival* *desinential paradigm*?

Question Two. *Why is the Orphan Accusative form the same as that of the genitive, rather than, say, that of the dative or instrumental case*?

Question Three. *Why is it all constituents with adjectival endings, rather than just quantifiers or demonstratives or just those adjectives that refer to inherent rather than transitory properties, that have a special form for the Orphan Accusative*?

Question Four. *Why is it that masculines have a special form for the Orphan Accusative and feminines do not*?

Question Five. *Why is it the accusative that has a special orphan form, rather than some other case*?

Question Six. *Why is there a special form in the absence of the head noun, rather than in some other environment, such as the presence of the head noun, or the absence of a quantifier, or when the head noun designates a stick-shaped or leafy object*?

The proposed Orphan Accusative Rule raises another question that derives not from its content but from its very existence. Why should there be such a rule in Slovenian, but not in, say, Latin or Finnish? There is a more general formulation of this question:

Question Seven. *What is particular to Slovenian in the Orphan Accusative phenomenon, and what is more general*?

Viewing this situation from the point of view of linguistic theory, we see that a linguistic theory that allows the formulation of language-particular rules such as the Orphan Accusative Rule in the grammars of particular languages thereby makes a wide range of similar rules available to particular grammars. In the absence of additional constraints on the form and content of syntactic rules, a theory that allows the Orphan Accusative Rule would also allow, among others, the following rules:

(22) A feminine plural accusative adjective takes the masculine singular instrumental ending if the head noun is not deleted (and the usual feminine plural accusative ending otherwise).

(23) A masculine singular dative adjective takes the ending -buni if there is only one adjective modifying the head noun.

(24) A masculine plural genitive or feminine singular locative demonstrative takes the feminine singular genitive ending in the absence of the head noun.

But it seems that no language has rules like (22-24). A theory of language that permits such rules therefore fails to constrain sufficiently the notion 'possible syntactic rule' and therefore also fails to constrain sufficiently the notion 'possible human language'. Thus, a linguistic theory that allows rules like the Orphan Accusative Rule and (22-24) is inadequate as a theory of language.

The immediate problem here is to replace the Orphan Accusative Rule by another analysis that will be able to answer Questions One to Seven. The more difficult problem is to construct a theory of language that will provide a principled basis for choosing such rules over the Orphan Accusative Rule, and will thereby exclude rules like (22-24) in principle, thus constraining more stridently the notions 'possible syntactic rule' and 'possible human language'.

4. Identity of Sense Pronominalization in Slovenian.

The difference between Identity of Reference Pronominalization and Identity of Sense Pronominalization in English is illustrated by the contrast between the sentences

(25) Bill saw a blue car and Tom saw it too.

and

(26) Bill saw a blue car and Tom saw one too.

The pronoun *it* in (25) indicates Identity of Reference (coreferentiality) between the car that Tom saw and the car that Bill saw, while the anaphoric element *one* in (26) indicates only that Tom saw a blue car -- though not necessarily the same one that Bill saw.

The Slovenian sentence

(27) Stane je videl plav avto in Tone ga je videl tudi.
saw blue car and it saw too
'Stane saw a blue car and Tone saw it too.'
'Stane saw a blue car and Tone saw one too.'

is ambiguous. The pronoun *ga* in (27) could be used either for Identity of Reference Pronominalization or for Identity of Sense Pronominalization. The fact that definite pronouns like *ga* can be used for Identity of Sense Pronominalization can also be seen clearly in the fact that they appear in sentences in which Identity of Reference Pronominalization would be impossible.

(28) Stane ima pametnega otroka in Tone ga ima tudi.
has bright child and him has too
'Stane has a bright child and Tone has one too.'
(Literally: 'Stane has a bright child and Tone has him too.')

(29) Stane ima pametno ženo in Tone jo ima tudi.
has bright wife and her has too
'Stane has a bright wife, and Tone has one too.'
(Literally: 'Stane has a bright wife and Tone has her too.')

Slovenian thus appears not to distinguish between Identity of Sense Pronominalization and Identity of Reference Pronominalization in surface structure. When we look deeper, however, we find that the definite pronouns like ga and jo are used only in some cases of Identity of Sense Pronominalization, not in all.

(30) a. Stane ima staro rjavo hišo in Tone ima novo belo.
has old brown house and has new white
'Stane has an old brown house and Tone has a new white one.'

b. Stane ima staro rjavo hišo in Tone ima novo.
has old brown house and has new
'Stane has an old brown house and Tone has a new one.'[4]

c. Stane ima staro rjavo hišo in Tone jo ima tudi.
has old brown house and it has too
'Stane has an old brown house and Tone has one too.'

Comparing the anaphoric elements in (30) with their meanings, we find the following correspondences:

(31)

Example	*Anaphoric element*	*Meaning*
(30a)	novo belo	novo belo hišo
(30b)	novo	novo rjavo hišo[4]
(30c)	jo	staro rjavo hišo

A grammar of Slovenian must associate the anaphoric elements in (31) with their meanings.

In each case, that part of the meaning that is identical to something in the first conjunct is not present in the second conjunct in surface structure. It is only in the case where the entire noun phrase is identical to the antecedent that the pronoun *jo* appears. It is possible to account for these facts by postulating that Identity of Sense Pronominalization in Slovenian proceeds in two stages.

(32) *Identity of Sense Pronominalization*

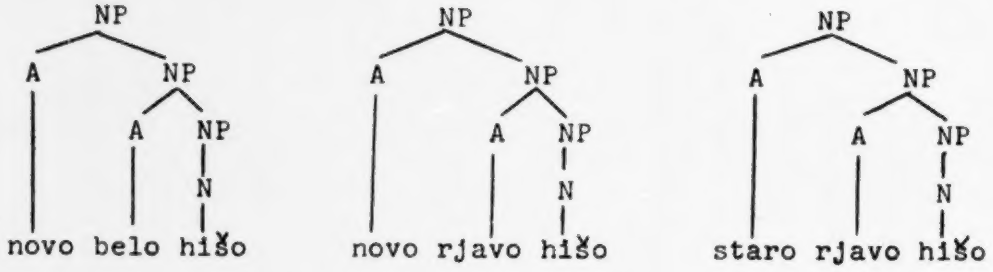
Stage One (Pronominalization): A noun phrase identical to an antecedent noun phrase is replaced by a pronoun.

Stage Two (Pronoun Deletion): The pronoun is deleted if it follows a modifier.

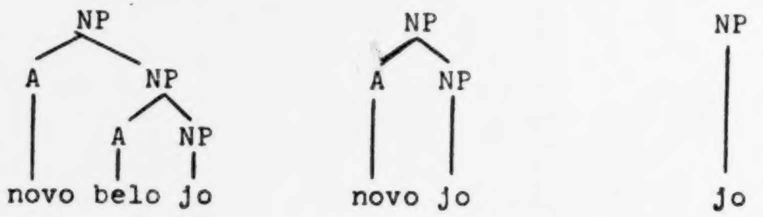
Under this analysis, the anaphoric elements in (30/31) would have the following derivations:[5]

Example: (30a) (30b) (30c)

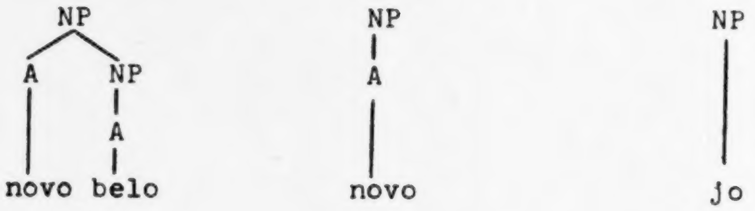
__Underlying Form:__


NP
A
NP
A
NP
N
novo belo hišo
NP
A
NP
A
NP
N
novo rjavo hišo
NP
A
NP
A
NP
N
staro rjavo hišo

__Stage One (Pronominalization):__


NP
A
NP
A
NP
novo belo jo
NP
A
NP
novo jo
NP
jo

__Stage Two (Pronoun Deletion):__


NP
A
NP
A
novo belo
NP
A
novo
NP
jo

We propose, then, that Identity of Sense Pronominalization in Slovenian proceeds in two stages -- Pronominalization of the identical noun phrase, followed by Pronoun Deletion if a modifier precedes.[6] This analysis is motivated by the necessity of matching the anaphoric elements in (30/31) with their meanings. It does that by taking structures that directly represent their meanings to be their underlying forms, and by deriving their surface forms from them in two stages.

There are (at least) two other possibilities. One might take either the structures produced by our Stage One (Pronominalization), to be the underlying forms, or else the structures produced by our Stage Two (Pronoun Deletion) to be the underlying forms. In the latter case, the underlying and surface forms of the anaphoric elements in (30/31) would be identical. Under either of these approaches, some other kind of mechanism would be necessary to associate the underlying forms with their meanings. We will not argue either for or against the proposal that our Stage One be taken as the underlying forms, since that issue is not relevant to the topic of this paper. The evidence presented in the rest of this paper, however, provides very strong arguments against taking our Stage Two (the surface forms) as the underlying anaphoric elements forms of the anaphoric elements in (30/31).

5. The Concord Hypothesis

Consider Identity of Sense Pronominalization with an accusative masculine noun. From the structure underlying

(33) Stane ima rjav površnik in Tone ima tudi rjav površnik.
has brown overcoat and has also brown overcoat
'Stane has a brown overcoat and Tone has a brown overcoat too.'

Identity of Sense Pronominalization will produce

(34) Stane ima rjav površnik in Tone ga ima tudi.
has brown overcoat and it has too
'Stane has a brown overcoat and Tone has one too.'

From the structure underlying

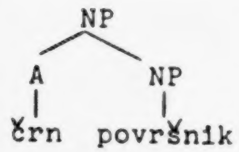
(35) Stane ima rjav površnik in Tone ima črn površnik.
has brown overcoat and has black overcoat
'Stane has a brown overcoat and Tone has a black overcoat.'

Identity of Sense Pronominalization will produce

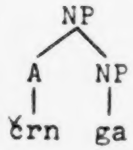
(36) Stane ima rjav površnik in Tone ima črnega.
has brown overcoat and has black
'Stane has a brown overcoat and Tone has a black one.'

In (36) we find the Orphan Accusative form črnega. The ordinary accusative črn would be ungrammatical in (36), as would the Orphan Accusative črnega in (35). This was also the case with the examples considered in §2, where the Orphan Accusative appeared in the absence of the head noun. In (36), too, the head noun is absent, and we find the Orphan Accusative. However, we now know more about the derivation of (36). In particular, we know that Identity of Sense Pronominalization takes place in two stages -- Pronominalization followed by Pronoun Deletion. The final noun phrase in (36) would therefore have to go through the following derivation:

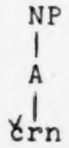
(37) *Underlying Form*:



Stage One (Pronominalization):



Stage Two (Pronoun Deletion):



If nothing else happens in the course of the derivation, we will get the incorrect *črn* instead of the Orphan Accusative form *črnega*.

We propose that the Orphan Accusative is produced by the rule of *Concord within the Noun Phrase* (henceforth simply 'Concord'), applying between Stage One and Stage Two of the derivation. The rule of Concord is needed in the grammar anyway, completely independently of the Orphan Accusative. Concord makes adjectives agree with the head noun in gender, number, and case, as the following examples illustrate:

(38) a. Stane ima star*o* rjav*o* hiš*o*.

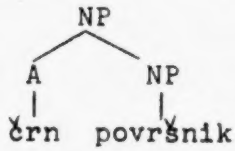
'Stane has an old brown house.'

b. Stane ima star rjav površnik.

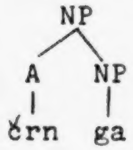
'Stane has an old brown overcoat.'

We propose that the Orphan Accusative arises from the application of this rule of Concord at the stage of derivations at which the underlying head noun has been replaced by a pronoun. Agreement of the adjective črn in (37) with the pronominal head ga at this stage of derivations will produce the Orphan Accusative form črnega. The mechanism by which Concord will produce črnega will be discussed in §6. In this section, we will limit ourselves to testing the consequences of the hypothesis that the Orphan Accusative is due to Concord, rather than to a special Orphan Accusative rule. After Concord has applied, the pronoun ga that causes Concord to produce the orphan form črnega will then be deleted by Pronoun Deletion, leaving the Orphan Accusative form črnega. The entire derivation will therefore be as follows:

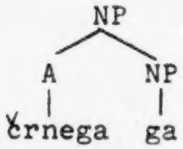
(39) Underlying Form:



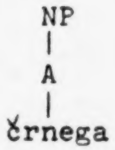
Pronominalization:



Concord:



Pronoun Deletion:



The most obvious consequence of using the rule of Concord to produce the Orphan Accusative in Slovenian is that the grammar is thereby simplified. Now there will be no special Orphan Accusative Rule at all, since the rule of Concord will do its work.

More importantly, the proposal to use the mechanism of Concord to produce the Orphan Accusative immediately makes two clear predictions that the Orphan Accusative Rule (21) does not make.

The Orphan Accusative Rule (21) is silent on the question of what happens to a masculine accusative noun phrase which contains more than one adjective if the head noun is absent. A priori, there are many possibilities, for example:

(40) a. Only the first adjective goes into the Orphan Accusative.
b. Only the last adjective goes into the Orphan Accusative.
c. Every second adjective goes into the Orphan Accusative.
d. Every third adjective goes into the Orphan Accusative.
e. All adjectives go into the Orphan Accusative.

and many others. The Concord hypothesis, however, predicts that all modifying adjectives will go into the Orphan Accusative, since they all agree with the head noun when there is one. And this prediction is correct. Thus, alongside (36) we find

(41) Stane ima star rjav površnik in Tone ima novega črnega.
has old brown overcoat and has new black
'Stane has an old brown overcoat and Tone has a new black one.'

with both adjectives novega and črnega in the Orphan Accusative.

Similarly, in answer to the question

(42) Koliko površnikov ima Stane?
how many overcoats has
'How many overcoats does Stane have?

one can answer

(43) Samo enega starega rjavega.
only one old brown
'Only one old brown one.'

in which the adjectives starega and rjavega, as well as the number 'one' enega, which always undergoes Concord to agree with the head noun, are in the Orphan Accusative.

Now consider how the facts illustrated by (42) and (43) might be incorporated into a rule like the Orphan Accusative Rule. It would be necessary to say that all adjectives in the noun phrase whose head is absent go into the Orphan Accusative. But closer inspection reveals that this is not the case. To see this, consider a sentence like

(44) Videl sem velik zemljevid, obsegajoč hrvatski okraj,
saw large map comprising Croatian district

in majhen zemljevid, obsegajoč slovenski okraj.
and small map comprising Slovene district

'I saw a large map, comprising a Croatian district, and a small map, comprising a Slovene district.'

The head noun in the second conjunct is zemljevid 'map', and its modifiers are majhen 'small', on the one hand, and obsegajoč slovenski okraj 'comprising a Slovene district', on the other.

What will happen if the head noun zemljevid is deleted? Will all masculine singular accusative adjectives go into the Orphan Accusative? If they are all put in the Orphan Accusative, the result is ungrammatical:

(45)*Videl sem velik zemljevid, obsegajoč hrvatski okraj, in
majhnega, obsegajočega slovenskega okraj.

The correct version of (44) with the second instance of zemljevid deleted is rather:

(46) Videl sem velik zemljevid, obsegajoč hrvatski okraj,
saw large map comprising Croatian district

in majhnega, obsegajočega slovenski okraj.
and small comprising Slovene district

'I saw a large map, comprising a Croatian district, and a small one, comprising a Slovene district.'

Crucially, the masculine singular accusative adjective slovenski in (46) does not go into the Orphan Accusative. And this fact is predicted automatically by the Concord hypothesis, for slovenski in (44) and (46) is not agreeing with zemljevid, but rather with its own head noun okraj 'district'. This can readily be seen by substituting the neuter noun mesto 'town' for okraj in (46):

(47) Videl sem velik zemljevid, obsegajoč hrvatski okraj, in
majhnega, obsegajočega slovensko mesto.

'I saw a large map, comprising a Croatian district, and a small one, comprising a Slovene town.'

Now we have the neuter form slovensko, in agreement with mesto. If the grammar contained a special Orphan Accusative Rule, the rule would have to state a condition something like:

(48) In the absence of the head noun, all modifying masculine singular accusative adjectives go into the Orphan Accusative, but only those modifying adjectives that modify the deleted head noun go into the Orphan Accusative.

This statement, however, duplicates the rule of Concord itself, and immediately leads to the question:

Question Eight. Why do all adjectives (rather than just one, or some) that modify the absent head go into the Orphan Accusative and why does this happen only to adjectives that actually modify the absent head?

By using the mechanism of Concord to produce the Orphan Accusative, the Concord hypothesis answers this question and provides an explanation of why it is all those adjectives that modify the deleted head, and only those, that go into the Orphan Accusative.

Using the mechanism of Concord to produce the Orphan Accusative also gives automatic answers to two of the questions already posed in §3.

Question Three. Why is it all constituents with adjectival endings, rather than just quantifiers or demonstratives or just those adjectives that refer to inherent rather than transitory properties, that have a special form for the Orphan Accusative? It is all constituents with adjectival endings that have a special form for the Orphan Accusative because they are the constituents that agree with the head noun. A rewrite rule such as (21) is perfectly capable of applying to any subclass of adjectives or to any class of constituents other than adjectives, but using Concord to produce the Orphan Accusative automatically predicts that whatever class of constituents agrees with the head noun will be the class that has a special form for the Orphan Accusative.

Question Six. Why is there a special form in the absence of the head noun, rather than in some other environment, such as the presence of the head noun, or the absence of a quantifier, or when the head noun is a stick-shaped or leafy object? Since the Orphan Accusative is produced by Concord with a pronominal head, it can arise only in an environment in which the head is pronominal. We have seen that Pronominalization is an intermediate stage in the process of deletion. Therefore we find the Orphan Accusative when the head noun has been deleted. There can be no Orphan Accusative in the presence of the head noun, since then there is no pronominal head, and it is Concord with a pronominal head that produces the Orphan Accusative. For the same reason, there can be no Orphan Accusative in any other environment in which there is no pronominal head.

In answering Questions Three and Six, the proposal to use the mechanism of Concord to produce the Orphan Accusative provides an explanation of these facts.

6. The Mechanics of Concord

We will now attempt to make precise how the rule of Concord produces the Orphan Accusative.

The conception of Pronominalization that was dominant in generative grammar in the mid-1960's received its clearest formulation in Postal (1966). Under this theory, constituents were conceived of as bundles of features, and Pronominalization consisted of adding the feature [+Pro] to the noun undergoing Pronominalization. The presence of the feature [+Pro] would then affect the operation of late *spellout rules* or *segmentalization rules* that gave morphemes their phonological shape.

It is easy to see that this conception of Pronominalization does not make it possible to use the rule of Concord to produce the Orphan Accusative in Slovenian in a natural way. In the sentence

(49) Tone ima črn površnik.
has black overcoat
'Tone has a black overcoat.'

the noun *površnik* would be represented as a bundle of features:

(50)
$$\begin{bmatrix}\text{+Noun}\\ \text{-Animate}\\ \text{+Singular}\\ \text{+Accusative}\\ \text{+Masculine}\\ \vdots\end{bmatrix}$$

The rule of Concord would give the modifying adjective the features:

(51)
$$\begin{bmatrix}\text{-Animate}\\ \text{+Singular}\\ \text{+Accusative}\\ \text{+Masculine}\end{bmatrix}$$

and then spellout rules or segmentalization rules that actualize these features as phonological shape would spell out the feature bundle (51) on the adjective *črn* with the zero ending, yielding the correct form črn in (49).

If Pronominalization merely adds the feature [+Pro] to a feature bundle, how will the Orphan Accusative form črnega be produced in

(52) Stane ima rjav površnik in Tone ima črnega.
has brown overcoat and has black
'Stane has a brown overcoat and Tone has a black one.'

The structure underlying (52) will have (49) as its second conjunct. površnik will be marked as in (50), and the addition of the feature [+Pro] to this feature bundle will in no way affect the adjective's being given the features in (51) by the rule of Concord. When the spellout or segmentalization rules apply, the adjective will get the zero ending, as it did in (49), and we will get the ungrammatical sentence

(53)*Stane ima rjav površnik in Tone ima črn.

instead of the grammatical (52).

In order to use the mechanism of Concord to account for the Orphan Accusative, we propose that the pronoun ga be marked in the lexicon with the feature [+Animate], and that the pronoun ga be actually present in post-Pronominalization trees. If Pronominalization is a transformation, it consists not of adding the feature [+Pro], but rather of replacing a noun phrase by a pronoun, with its own feature bundle. If, on the other hand, Pronominalization is not a transformation, but pronouns are already present in underlying structures, then ga will of course already be present in trees. Our solution requires only that the pronoun ga be actually present in post-Pronominalization trees; it does not matter whether it was already present in underlying structures or produced by a transformation. If the pronoun ga is lexically marked as [+Animate] and is actually present in post-Pronominalization trees, then the rule of Concord will make adjectives that modify ga [+Animate] as well. These adjectives will therefore not have the features in (51), but rather:

(54) $\begin{bmatrix} \text{+Animate} \\ \text{+Singular} \\ \text{+Accusative} \\ \text{+Masculine} \end{bmatrix}$

Then, the application of the Accusative Prediction Rule to adjectives with the features in (54) will give them the accusative ending -ega, which is their genitive ending as well. If the pronoun ga is inherently [+Animate] and is actually present in post-Pronominalization trees, then, application of Concord at a stage of derivations after Pronominalization and before Pronoun Deletion will produce the grammatical sentence (52) and prevent *(53).

With respect to our proposal that the pronoun ga is idiosyncratically marked [+Animate], it should be noted that neither this feature nor the device of marking lexical items idiosyncratically for this feature is ad hoc. The feature [±Animate] exists in Slovenian grammar independently of the Orphan Accusative, since the Accusative Prediction Rule makes crucial use of it. And although the feature [±Animate] on a particular noun is, in the overwhelming majority of cases, predictable from the semantic representation of the noun in question, there are a few nouns which idiosyncratically act as animates with respect to the Accusative

Prediction Rule, even though they are semantically inanimate. Thus, rak 'cancer' and as 'ace', for example, have their accusative like their genitive (raka and asa, respectively) instead of like their nominative (rak and as), even though they are semantically inanimate. They must therefore be marked as inherently [+Animate] in the lexicon.[7] Thus the device of marking particular lexical items [+Animate] is needed in the grammar of Slovenian anyway, independently of the Orphan Accusative. In proposing that the pronoun ga be lexically marked [+Animate], then, we are only proposing another use for a device that is needed anyway.

Our proposal to account for the Orphan Accusative by marking ga as [+Animate] in the lexicon and let the rule of Concord and the Accusative Prediction Rule produce the Orphan Accusative constitutes our answer to Question Seven in §3:

What is particular to Slovenian in the Orphan Accusative phenomenon, and what is more general? What is particular to Slovenian is the fact that the pronoun ga is inherently marked with the feature [+Animate].[7] Everything else is perfectly general. The Accusative Prediction Rule, the rule of Concord, and the breakdown of Identity of Sense Pronominalization into Pronominalization and Pronoun Deletion are needed in the grammar of Slovenian anyway, independently of the Orphan Accusative. Furthermore, our solution to the problem of the Orphan Accusative does not involve placing any restrictions on the operation of transformations. Rules apply in perfectly general fashion. With the pronoun ga marked [+Animate], the application of independently motivated rules automatically produces the Orphan Accusative. Our solution is therefore in accord with the view that the rules of a language express regularities, while idiosyncrasies are to be found in the lexicon.

We will now turn to the remaining questions posed in §3 and show how this solution automatically answers them. In so doing, it provides an explanation of the facts in each case.

Question One. Why is it that the special form found in the Orphan Accusative is not an arbitrary ending such as -uruburu or -gnuf, but rather an ending that exists elsewhere in the adjectival desinential paradigm? A rewrite rule such as (21) would be capable of producing an arbitrary ending such as -uruburu or -gnuf for the Orphan Accusative. Under our hypothesis, however, the Orphan Accusative arises from Concord with a pronoun marked [+Animate]. The Orphan Accusative form is therefore the usual form of adjectives that agree with an animate accusative head.

Question Two. Why is the Orphan Accusative form the same as that of the genitive, rather than, say, that of the dative or instrumental case? The form for animate masculines in the accusative singular is like the genitive. Since the Orphan Accusative is produced by agreement with an animate head, its ending is like the genitive.

Question Four. Why is it that masculines have a special form for the Orphan Accusative and feminines do not? Under our hypothesis, the Orphan Accusative is the result of the Accusative Prediction Rule applying to adjectives that have acquired the feature [+Animate] by Concord. Since the Accusative Prediction Rule applies to masculines but not to feminines, it follows automatically that masculines have a special Orphan Accusative form and feminines do not.

Question Five. Why is it the accusative that has a special orphan form, rather than some other case? Since the Orphan Accusative results from the application of the Accusative Prediction Rule to [+Animate] adjectives, and since the Accusative Prediction Rule specifies accusative endings rather than those of some other case, it is the accusative and not some other case that has a special orphan form.

The proposal to mark the pronoun ga with the feature [+Animate] and to let the rule of Concord distribute this feature to adjectives and the Accusative Prediction Rule spell it out as a genitive ending thus automatically answers the remaining questions posed in §3. It thus gives an explanation of why Slovenian has the particular distribution of the Orphan Accusative that it has, rather than some other imaginable distribution.

7. The Orphan Accusative in the Plural and Dual.

While the accusative singular has a special orphan form that is like the genitive, as we have seen, this is not the case in the plural and the dual. Thus, in answer to the question

(55) Katere površnike hočete?

'Which overcoats do you want?

one could answer either

(56) Hočem navadne površnike.

'I want ordinary overcoats.'

or

(57) Hočem navadne.

'I want ordinary ones.'

Use of the genitive form of the adjective would be ungrammatical:

(58)*Hočem navadnih.

Similarly, one could answer (55) with either

(59) Navadne površnike.

'Ordinary overcoats.'

or

(60) Navadne.

'Ordinary ones.'

but not with

(61)*Navadnih.

It is the same in the dual. In answer to

(62) Katera površnika hočete?

'Which (two) overcoats do you want?'

one can answer

(63) a. Hočem navadna površnika.

'I want (two) ordinary overcoats.'

b. Hočem navadna.

'I want (two) ordinary ones.'

c. Navadna.

'(Two) ordinary ones.'

but not

(64) a. *Hočem navadnih.

b. *Navadnih.

with the genitive form navadnih instead of the ordinary accusative navadna.

This immediately raises a question:

Question Nine. Why is the Orphan Accusative found only in the singular, and not in the plural or dual?

A rule such as the Orphan Accusative Rule (21) would be able to state this fact simply by adding the specification [+Singular] to the left-hand side of the arrow. But the hypothesis proposed here provides an explanation. It says that the Orphan Accusative arises only as the result of the confluence of two factors: a [+Animate] marking on the pronominal head when Concord applies, and the application of the Accusative Prediction Rule in giving animate accusatives genitive endings. Our hypothesis therefore predicts that Orphan Accusatives will occur only in those morphological categories in which the Accusative Prediction Rule applies. If there are morphological categories in which the Accusative Prediction Rule does not apply, then, even if a particular pronoun is marked [+Animate], there will be no Orphan Accusative, because the Orphan Accusative is the result of the Accusative Prediction Rule spelling out a [+Animate, +Accusative] ending as a genitive.

In the plural, we find the following paradigms:[8]

	Animate Plural	Inanimate Plural
NOMINATIVE	navadni slovenci	navadni površniki
ACCUSATIVE	navadne slovence	navadne površnike
GENITIVE	navadnih slovencev	navadnih površnikov
	'ordinary Slovenes'	'ordinary overcoats'

These paradigms show that the Accusative Prediction Rule does not apply in the plural. Thus, even if the accusative plural pronoun jih is marked [+Animate], the [+Animate] marking placed on modifying adjectives by the Concord rule will not be spelled out as a genitive ending by the Accusative Prediction Rule in the plural. Our hypothesis thus automatically predicts that there can be no Orphan Accusative in the plural.

In the dual, the paradigms are as follows:

	Animate Dual	Inanimate Dual
NOMINATIVE	navadna slovenca	navadna površnika
ACCUSATIVE	navadna slovenca	navadna površnika
GENITIVE	navadnih slovencev	navadnih površnikov
	'ordinary Slovenes'	'ordinary overcoats'

These paradigms show that the Accusative Prediction Rule must be reformulated so that (1a) applies only in the masculine singular. The exact reformulation does not concern us here. What is important here are the consequences of this reformulation for the Orphan Accusative. Under our hypothesis, the Orphan Accusative results whenever the Accusative Prediction Rule spells out animate accusative endings as genitive. Since the Accusative Prediction Rule does not do this in the dual, it follows automatically that, even if the accusative dual pronoun ju is marked [+Animate], there will be no Orphan Accusative in the dual.

Note that, with the limitations that have been observed on the domain of the Accusative Prediction Rule, it is possible that the [+Animate] marking is not confined to the masculine singular accusative pronoun ga, but is to be found on all pronouns in Slovenian;[9] it would then be only the restricted domain of the Accusative Prediction Rule that is responsible for the restricted distribution of a special Orphan Accusative form. In §8, evidence will be presented to show that this is in fact the case.

In §7, it has been shown that, since our hypothesis makes use of the Accusative Prediction Rule to produce the Orphan Accusative by the same mechanism that makes the animate accusative like the genitive, and since the Accusative Prediction Rule does not do this in the plural and dual, the hypothesis automatically answers Question Nine, providing an explanation of why there is no Orphan Accusative in the plural and dual.

8. The Orphan Accusative in the Neuter Singular.

Let us now turn to the question of the Orphan Accusative in the neuter singular, which we will illustrate with the noun proso 'millet'. In answer to the question

(65) Katero proso hočete?

'Which millet do you want?'

one can answer

(66) Hočem navadno proso.
'I want ordinary millet.'

or

(67) Navadno proso.
'Ordinary millet.'

It is also possible to delete the head noun *proso*, but Slovenes tend to be hesitant to do this if the head is neuter. The reason for the hesitation is an uncertainty about what to do with the resulting orphaned adjective. Should one say

(68) Hočem navadno.
'I want ordinary.'

or

(69) Navadno.
'Ordinary.'

with the ordinary accusative, or should one say

(70) Hočem navadnega.
'I want ordinary.'

or

(71) Navadnega.
'Ordinary.'

with the Orphan Accusative? The situation here has two interesting properties: first, speakers are uncertain as to whether to use (68) and (69) or (70) and (71); second, they tend to accept *all four* possibilities as grammatical. The situation in the neuter singular therefore differs from that in the masculine singular. In the masculine, Slovenes are not hesitant and uncertain, and the

Orphan Accusative is obligatory and not optional. The question

(72) Kateri ječmen hočete?

'Which barley do you want?'

can be answered by

(73) a. Hočem navadnega.

'I want ordinary.'

b. Navadnega.

'Ordinary.'

with the Orphan Accusative, but not by

(74) a. *Hočem navaden.

b. *Navaden.

with the ordinary accusative. The difference between the masculine and neuter singular must be explained:

Question Ten. Why are speakers uncertain as to whether or not to use the Orphan Accusative in the neuter singular, and why is its use there optional?

If the neuter singular accusative pronoun ga is marked [+Animate], our hypothesis answers Question Ten automatically, without any additional machinery.

The reason for this is that, under our hypothesis, if the pronoun is marked [+Animate], an Orphan Accusative form will show up precisely where the Accusative Prediction Rule gives [+Animate, +Accusative] endings the same shape as genitives. In order to see what our hypothesis predicts about the neuter singular, it is therefore necessary to see whether or not the Accusative Prediction Rule does this in the neuter singular. Inanimate neuter nouns have their accusative like their nominative, but it is the animate neuter nouns that are crucial here. There are only a handful of them, and, when asked about their form in the accusative, Slovenes tend to be uncertain as to whether the accusative is like the nominative or like the genitive. After thinking it over, most speakers asked tended to accept both forms as grammatical. Thus we find the following paradigms with animate neuter singular nouns:

NOMINATIVE	navadno dekle	navadno ščene
ACCUSATIVE	navadno dekle/navadnega dekleta	navadno ščene/navadnega ščeneta
GENITIVE	navadnega dekleta	navadnega ščeneta
	'ordinary girl'	'ordinary puppy'

Regardless of what the precise reason for this may be,[10] the fact is that Slovenes tend to be uncertain as to whether or not the accusative of animate neuter singular nouns is like the nominative or like the genitive, and they tend to accept both forms as grammatical.

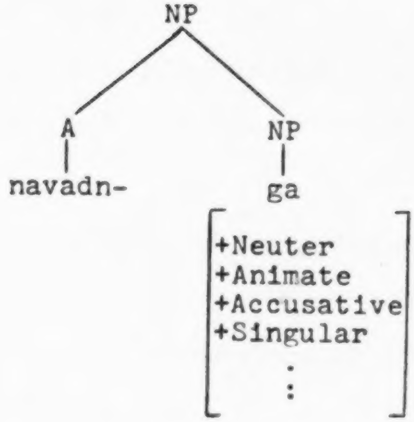
The feeling of uncertainty and the acceptance of both possibilities is precisely what characterizes the situation with respect to the Orphan Accusative in the neuter singular. If the neuter singular accusative pronoun ga is marked [+Animate], both the uncertainty and the acceptability of both forms will follow automatically from our hypothesis. Consider the derivation of an answer to the question

(75) Katero proso hočete?

'Which millet do you want?'

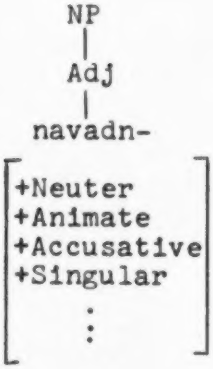
At the relevant stage of derivations, the answer will have the form:

(76)


NP
A
NP
navadn-
ga
+Neuter
+Animate
+Accusative
+Singular

Concord will place the features [+Neuter, +Animate, +Accusative, +Singular] on the adjective, and after Pronoun Deletion we will be left with:

(77)


NP
Adj
navadn-
+Neuter
+Animate
+Accusative
+Singular

The Accusative Prediction Rule should now apply to (77). But, as we have seen from the paradigms of *dekle* and *ščene*, people are uncertain as to whether or not the rule makes the accusative of animate neuters the same as the genitive, and tend to accept both possibilities. If it does, (77) will be actualized as

(78) Navadnega.

and if it does not, we will get

(79) Navadno.

Speakers' uncertainty as to whether or not the accusative is like the genitive in the paradigms of *dekle* and *ščene* is automatically translated by our hypothesis into their uncertainty as to whether to answer (75) by (78) or by (79), and their acceptance of both *navadno dekle* and *navadnega dekleta* is translated into their acceptance of both (78) and (79). The initially strange behavior of neuter singulars with respect to the Orphan Accusative thus follows as an automatic consequence of their strange behavior with respect to the Accusative Prediction Rule.

It should be noted that the masculine singular accusative pronoun *ga* and the neuter singular accusative pronoun *ga* are not the same entity, as might be thought, but rather are distinct. The evidence for this comes from the derivation of (78) and (79) from (77). The adjective in (77) must be marked [+Neuter] in order for it to evoke uncertainty and optionality with respect to the Accusative Prediction Rule; if it were the same as the masculine pronoun, there would be no uncertainty, and *navadnega* would be the only possible grammatical output. The conclusion is inescapable that the masculine accusative pronoun *ga* and the neuter accusative pronoun *ga* are distinct.

Both pronouns, however, must be marked [+Animate] if the appearance of the Orphan Accusative in both genders is to be accounted for. It would be strange enough for two pronouns to have *ad hoc* feature markings, but even stranger for them both to have the *same* ad hoc markings. A grammar in which the animacy of the masculine pronoun and that of the neuter pronoun are two separate and unrelated facts is missing a generalization. However, there is a way that the relevant generalization can be captured. That is to say that all pronouns in Slovenian are marked [+Animate].[11] Since the only place in Slovenian where animacy shows up syntactically is where the Accusative Prediction Rule makes the accusative of animates like their genitive, and since, as we have seen, this rule applies only in the masculine singular and, with uncertainty and optionality, in the neuter singular, the animacy markings on the other pronouns will cause no difficulty; there are no other syntactic rules that refer to animacy.

Although Slovenian pronouns must be marked [+Animate] to account for the appearance of the Orphan Accusative, they can still refer to inanimate objects, as in

(80) Kar zadeva površnik, sem ga videl včeraj.

'As far as the overcoat is concerned, I saw it yesterday.'

in which the pronoun ga refers to the inanimate površnik 'overcoat'. Yet, it is only the 'weak' or 'clitic' forms of the pronouns that can refer to inanimate objects; the corresponding 'strong' forms can not. Thus, in the sentence

(81) Včeraj sem videl samo njega.

'Yesterday I saw only him.'

in which we have the strong form njega instead of the clitic ga, njega can only have animate reference. Note, in this connection, that this is true not only of the strong form njega, which corresponds to the clitic ga, but also of all other strong form pronouns in Slovenian. Thus, the feminine singular strong form njo in

(82) Včeraj sem videl samo njo.

'Yesterday I saw only her.'

also can only have animate reference. This fact is of interest in connection with our conclusion that it is not just the masculine and neuter accusative singular pronouns, but all pronouns in Slovenian, that are marked for animacy. The [+Animate] marking that triggers the Orphan Accusative is no doubt related to the fact that the strong forms of the pronouns can only have animate reference. It is not clear, however, how this can be captured in the present theory of grammar. The question of how this is to be incorporated in a grammar and the concomitant question of why pronouns that are marked [+Animate] can refer to inanimate objects must be left for future research.

In §8 it has been shown how the status of the Orphan Accusative in the neuter singular follows automatically from the status of the Accusative Prediction Rule in the neuter singular. It followed from this that the neuter pronoun ga is distinct from the masculine pronoun ga, but is also marked [+Animate]. We concluded that the correct generalization is that all Slovenian pronouns are marked [+Animate].

9. Chopping Rules.

The Orphan Accusative Rule (21) stated that the Orphan Accusative arises 'in the absence of the head noun'. But it did not specify the notion 'absence' more precisely. This notion might be defined in any of three ways:

(83) a. A constituent is 'absent' if it has been deleted.

b. A constituent is 'absent' if it has been moved away by a chopping rule in the sense of Ross (1967).

c. A constituent is 'absent' if it has either been deleted or else moved away by a chopping rule.

A priori, there is no reason to choose any one of these over the other two. The hypothesis proposed here, on the other hand, states that the Orphan Accusative arises through Concord with a pronominal head, and it was shown that deletion passes through a pronominal stage. The hypothesis thus predicts that (83b) and (83c) are not possible definitions of 'absence' of the head, i.e., that if a constituent is moved away by a chopping rule, its modifiers will not go into the Orphan Accusative. It remains to test this prediction.

Under almost all circumstances, chopping rules do not move a head away from under its modifiers. They can, however, if the modifier is ves 'all'. Thus, alongside the sentence

(84) Želel sem pojesti ves riž.
wanted eat up all rice
'I wanted to eat up all the rice.'

we also find

(85) Riž sem želel pojesti ves.
rice wanted eat up all
'The rice I wanted to eat up all of.'

in which riž has been preposed. Crucially, the modifier ves has remained in the ordinary accusative form. Putting it in the Orphan

Accusative results in an ungrammatical sentence:

(86) *Riž sem želel pojesti vsega.

Given that the preposing rule is a chopping rule, this is exactly what our hypothesis predicts.

10. Direct Evidence for the Stage of Derivations at which Concord Applies.

All of the questions concerning the distribution of the Orphan Accusative have now been answered. This has been accomplished without adding any new rules or apparatus to the grammar -- merely by marking pronouns as [+Animate].

The stage of derivations at which noun phrases consist of modifiers and a pronominal head (e.g., črn ga) is crucial to the hypothesis, since it is Concord at this stage of derivations that produces črnega ga, which Pronoun Deletion then reduces to črnega, the Orphan Accusative form that appears in surface structure. Since Pronoun Deletion is obligatory if the pronoun is preceded by a modifier, we have not been able to give direct evidence for this stage of derivations. The evidence for its existence has been indirect: postulating it has enabled us to explain the distribution of the Orphan Accusative.

It would be possible to find direct evidence for the stage of derivations at which Concord applies if there is a rule which separates the pronominal head from its modifiers after Concord has applied but before the application of Pronoun Deletion. Under such circumstances, if our hypothesis is correct, a modifier that agrees with the masculine singular accusative pronominal head ga should have the -ega ending characteristic of the Orphan Accusative, with the head ga that triggered Concord still present in surface structure. Two additional properties of Slovenian -- Clitic Movement and the behavior of ves 'all' -- make it possible to find the relevant evidence.

In Slovenian, clitic pronouns are obligatorily moved to second position in the clause.[12] Thus, while 'I wanted to find Jože' is expressed by the sentence

(87) Želel sem najti Jožeta.
wanted find
'I wanted to find Jože.'

the sentence 'I wanted to find him' is expressed by

(88) Želel sem ga najti.
wanted him find
'I wanted to find him.'

in which the clitic pronoun ga has been moved from the position of Jožeta in (87) to second position, where it joins the auxiliary verb sem, the other clitic.[13]

Ves 'all' is an unusual modifier in that it allows a clitic pronoun head of its noun phrase to move out to second position in the sentence. Thus, while 'I wanted to eat up all the buckwheat' would be expressed as

(89) Želel sem pojesti vso ajdo.
wanted eat up all buckwheat
'I wanted to eat up all the buckwheat.'

'I wanted to eat up all of it' would be rendered as

(90) Želel sem jo pojesti vso.

wanted it eat up all

'I wanted to eat up all of it.'

in which the clitic pronoun jo has joined sem in second position, leaving behind the modifier vso.[14]

Now consider an example in which the head is masculine:

(91) Želel sem pojesti ves riž.

wanted eat up all rice

'I wanted to eat up all the rice.'

To say 'I wanted to eat up all of it,' referring to riž 'rice', one would say

(92) Želel sem ga pojesti vsega.

wanted it eat up all

'I wanted to eat up all of it.'

The clitic pronoun ga has moved into second position, leaving behind the modifier vsega. The sentence

(93) *Želel sem ga pojesti ves.

in which the modifier has been left in the ordinary accusative, is ungrammatical. This is a remarkable fact. It shows that, completely independently of the Orphan Accusative, the pronoun ga must be marked in some way so that Concord will produce the genitive-like form vsega instead of the ordinary accusative ves, if (92) and *(93) are to be accounted for. Thus the device we have posited to account for the Orphan Accusative is needed in the grammar anyway. (92) and *(93) show clearly that when a modifier agrees with the pronominal head ga, the modifier takes on the -*ega* ending that shows up on the Orphan Accusative. This is striking confirmation of the hypothesis presented here.

11. *Predicate Attribute Agreement with a Pronominal Trigger*.

Implicit in our hypothesis that the Orphan Accusative is produced by Concord with a pronominal head is another prediction: any adjective that agrees with a pronoun will acquire the feature [+Animate], which, as a result of the Accusative Prediction Rule, will have the accusative ending characteristic of animates, *regardless of which agreement rule puts the [+Animate] feature on the adjective*. Our discussion so far has been confined to examples where the relevant agreement rule has been Concord (within the noun phrase), but there is another agreement rule which can be used to test this prediction -- the rule we will refer to as Predicate Attribute Agreement or simply Attribute Agreement. This is the rule that is responsible for the agreement of *pomazan* 'stained' in:

(94) Včeraj smo našli mizo pomazano s krvjo.
yesterday found table stained with blood
'Yesterday we found the table stained with blood.'

(95) Včeraj smo našli stol pomazan s krvjo.
chair
'Yesterday we found the chair stained with blood.'

In (94), *pomazano* is feminine singular accusative, in agreement with its antecedent *mizo*, while in (95), *pomazan* is masculine singular accusative, in agreement with its antecedent *stol*.[15]

What our hypothesis predicts is that if the antecedent of pomazan is the pronoun ga, it will have the animate accusative -ega ending that is characteristic of the Orphan Accusative. And this is exactly what happens:

(96) Včeraj smo ga našli pomazanega s krvjo.

'Yesterday we found it stained with blood.'

The ordinary masculine singular accusative form pomazan, which is grammatical in (95), would not be in (96):

(97) *Včeraj smo ga našli pomazan s krvjo.

This is further confirmation of our claim that it is agreement with the pronoun ga that produces the -ega ending of the Orphan Accusative. This phenomenon is now seen not to be confined to cases of Concord within the noun phrase; it extends to all sentences in which a modifier agrees with a pronoun. Thus, there can be no language (or dialect of Slovenian) in which adjectives have the Orphan Accusative found in Slovenian but do not assume the same endings under agreement with an overt pronoun. It is this claim that is supported by (96) and *(97).

Footnotes

* This work was supported in part by a grant from the National Institutes of Health (5T01 HD00111-07) and a grant from the National Institutes of Mental Health (5 P01 MH13390-05) to the Massachusetts Institute of Technology and a grant from the United States Office of Education (OEC-0-70-4986(823)) to the Language Research Foundation. The authors are indebted to Stephen Anderson, Wayles Browne, Jorge Hankamer, Paul Postal, and John Ross, who read an earlier version of the manuscript and suggested many improvements. Responsibility for errors rests with the authors.

1. Additional facts, such as the existence of other declensional classes and a definite (as opposed to indefinite) adjective ending, are not relevant to the subject of this paper and are therefore ignored here.

2. A few exceptions which have their accusative like their genitive even though they are inanimate are discussed briefly in §6.

3. Since animate masculine nouns have their accusative like their genitive anyway, they do not show the distinction between the Orphan Accusative and the ordinary accusative overtly.

4. The Slovenian example, like its English translation, is ambiguous: nov can mean either 'new brown house' or just 'new house'.

5. The trees given here as 'underlying forms' are not claimed to be the 'deepest' level of representation. Thus, if adjectives are derived from underlying relative clauses, the 'underlying forms' below are derived structures. The question of whether these underlying structures are deep or derived is not relevant here.

6. No explanation is offered here for why pronoun deletion depends on the presence of a modifier. This fact, too, requires explanation.

7. As will be seen in what follows, even this idiosyncrasy is related to a more general phenomenon.

8. The e in the -ev ending of slovencev is due to a phonological rule that is sensitive to the preceding consonant; there is no difference between the underlying forms of the animate and inanimate endings.

9. We are indebted to Wayles Browne for this suggestion. This is possible because, except for the Accusative Prediction Rule, there is no other way that animacy shows up syntactically in Slovenian.

10. Two facts of Slovenian morphology are relevant here. First, in the dual and plural, the accusative of neuters is like the nominative. Second, throughout the singular paradigm (except for the nominative case, which is the citation form from which gender is in general predictable, and the accusative, which is under discussion here), neuter endings are the same as masculine endings. In the neuter singular, then, the conflict is between the generalization that neuter endings are like masculine endings in the singular,

as a result of which the Accusative Prediction Rule (1) should apply in the neuter singular and yield *navadnega dekleta* and *navadnega ščeneta*, and the generalization that the accusative of neuters is like the nominative throughout the paradigm, which would yield *navadno dekle* and *navadno ščene* as accusative singular forms.

11. Once all pronouns are marked [+Animate], the argument given in §6 against a feature treatment of Pronominalization (as consisting of marking constituents with the feature [+Pro]) loses its force, as Charles Fillmore has pointed out to us, for it could then be argued that the way to mark pronouns for animacy is by means of a rule of the form: [+Pro] ⟶ [+Animate]. Since pronouns in many languages behave like animates, the ultimate resolution of this question will depend on how the animate-like behaviors of pronouns are to be captured in universal grammar.

12. This is actually an oversimplification, for the facts of clitic placement in Slovenian are more complex. For discussion of this question, see Toporišič (1970, 176ff.).

13. The two clitics *sem* and *ga*, as a group, are in second position. Clitic order constraints of the type discussed in Perlmutter (1971, Chapter Two) specify that within the clitic group *sem* precedes *ga*.

14. *jo* is the feminine singular accusative pronoun, and *vso* is the feminine singular accusative form of *ves* 'all'.

15. We will not be concerned here with the problem of specifying which noun phrase is the 'antecedent' of the predicate attribute. For an important study of this problem, see Andrews (1971).

References

Andrews, Avery (1971) "Case agreement of predicate modifiers in ancient Greek," Linguistic Inquiry 2, 127-151.

Perlmutter, David M. (1971) Deep and Surface Structure Constraints in Syntax, Holt, Rinehart and Winston, Inc., New York.

Postal, Paul M. (1966) "On so-called 'pronouns' in English," in Report of the Seventeenth Annual Round Table Meeting on Linguistics and Language Studies, Georgetown University Press, Washington. Reprinted in R. Jacobs and P. Rosenbaum (eds.) Readings in English Transformational Grammar, Blaisdell Publishing Co., Waltham, Mass., 1970, and in D. Reibel and S. Schane (eds.) Modern Studies in English, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1969.

Toporišič, Jože (1970) Slovenski knjižni jezik 4, Založba Obzorja, Maribor.

LANGUAGE-PARTICULAR RULES AND EXPLANATION IN SYNTAX

Part Two - Consequences for Linguistic Theory

David M. Perlmutter
and
Janez Orešnik

Part Two - Consequences For Linguistic Theory

11. On Deletion

11.1 A Pronominal Stage in Deletion

Postal (1970b) argues that the constituents deleted by Equi-NP Deletion in English are not full noun phrases, but rather pronouns, at the stage of derivations at which they are deleted. He suggests that this is but a particular consequence of a universal principle that deletion of a noun phrase subject to the existence of a coreferent noun phrase in the same structure always passes through a pronominal stage. We have shown here that in Slovenian, identity of sense deletion also goes through a pronominal stage. This suggests that the universal proposed by Postal may be but a special case of a more general principle: any deletion of a noun phrase under identity (whether identity of sense or identity of reference) to another noun phrase passes through a pronominal stage. It remains to be seen whether a universal principle as strong as this one survives empirical test.

11.2 The Domain of Deletion

It has been shown here that at some stage of derivations, certain noun phrases in Slovenian have a pronominal head that is subsequently deleted. Nothing has been said, however, as to whether these pronominal heads are already present in underlying structures or are produced transformationally. Whichever is the case, however, such noun phrases with pronominal heads are possible only if preceded by a coreferential noun phrase. Thus, a sentence like

(94) Ima samo enega starega rjavega.
has only one old brown
'He has only one old brown one.'

is not possible in isolation. However, it is natural as an answer to the question

(95) Koliko površnikov ima Stane?
'How many overcoats does Stane have?'

Thus the antecedent of a pronominal head can lie in a previous sentence in the discourse.

A question that has not been resolved here is the question of which of the two phenomena involved in the derivation of sentences like (94)--Pronominalization (whether or not it is a transformation) or Pronoun Deletion--is the one that requires the presence of an antecedent. It seems likely that the antecedent requirement can be placed on Pronominalization. This seems natural in view of the fact that identity-of-reference pronouns require antecedents too. If this is correct, the rule of Pronoun Deletion can be a rather local rule that can be formulated roughly as follows:

(96) Pronoun Deletion

$$[\mathrm{X\ Adj\ }\underset{\emptyset}{\underset{\Downarrow}{\mathrm{Pro}}}]_{NP}$$

The other possibility is that it is the deletion rule itself--and not Pronominalization--that requires an antecedent noun phrase.

The resolution of this issue must be left for future research.

11.3 Directionality of Deletion

Consider a sentence like

(97) Videl sem očiščen ječmen in navaden ječmen.
seen cleaned barley and ordinary barley
'I saw cleaned barley and ordinary barley.'

with a conjoined noun phrase in the accusative case. The conjoined noun phrase can be reduced in either of two ways. If the second instance of ječmen disappears we get

(98) Videl sem očiščen ječmen in navadnega.
'I saw cleaned barley and regular.'

with navadnega in the Orphan Accusative. Failure to use the Orphan Accusative in (98) results in ungrammaticality:

(99) *Videl sem očiščen ječmen in navaden.

This means that the object NP in (98) passes through the following stages in its derivation:

(100) a. očiščen ječmen in navaden ga
b. očiščen ječmen in navadnega ga
c. očiščen ječmen in navadnega

Since deletion passes through a pronominal stage, with Concord applying when the head is a pronoun, we can conclude that (98) arises through deletion of the second instance of ječmen.

It is also possible to reduce the object in (97) in such a way that the first instance of ječmen disappears. But then the facts are quite different. We get

(101) Videl sem očiščen in navaden ječmen.

'I saw cleaned and regular barley.'

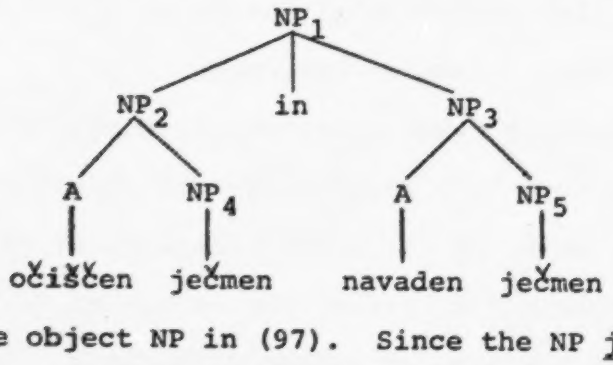
in which *očiščen* is *not* in the Orphan Accusative. Putting it in the Orphan Accusative results in ungrammaticality:

(102) *Videl sem očiščenega in navaden ječmen.

This is a striking fact. Why is there an asymmetry between (98) and (101) in the appearance of the Orphan Accusative?

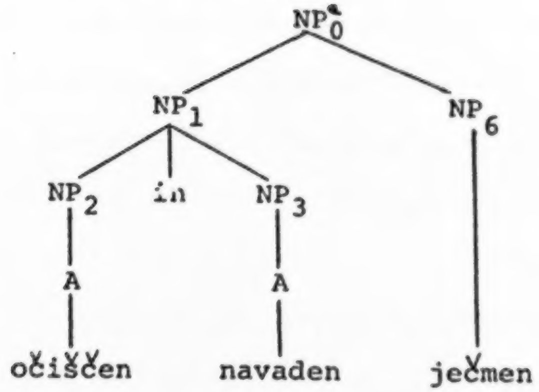
We have already seen that deletion passes through a pronominal stage, allowing Concord to produce the Orphan Accusative along the way. The fact that the Orphan Accusative does not appear in (101) therefore indicates that the reduction is not accomplished by deletion. To what, then, could this reduction be due? We propose that (101) is derived by means of the independently necessary rule of Right Node Raising. This rule, as formulated by Lakoff and Ross, takes any constituent that is an end-constituent in all conjuncts and removes it from all the conjuncts, Chomsky-adjoining a copy of it to the co-ordinating node. Right Node Raising will thus operate on the structure

(103)



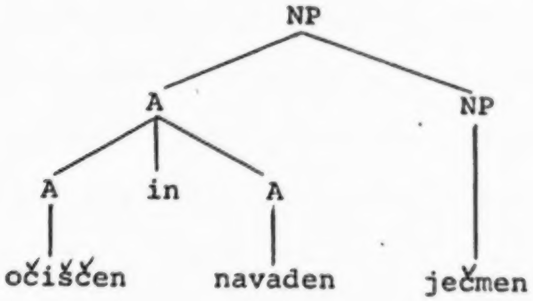
which is the object NP in (97). Since the NP *ječmen* is at the right end of both conjuncts, Right Node Raising will remove these two NPs from their conjuncts and Chomsky-adjoin a copy to the co-ordinating node (NP_1) to produce the structure

(104)



Under Lakoff and Ross's proposal, necessary principles of derived constituent structure will convert (104) into the structure

(105)



At no point in the derivation does *očiščen* modify a pronominal head. As a result, Concord will not put it in the Orphan Accusative, and we end up with the grammatical sentence (101).

The proposal that reduction of the first instance of *ječmen* in (97) is accomplished by Right Node Raising and that of the second by deletion thus accounts for the appearance of the Orphan Accusative in (98) and its ungrammaticality in *(102).

Since reduction--whether by deletion or by Right Node Raising--is optional, something additional is needed to prevent the wrong rule from applying in each case. Thus, if Right Node Raising produced a structure in which *ječmen* remained on the left and disappeared on the right, *(99) would be generated. But the way Right Node Raising is formulated, it cannot produce this output.

Thus there is no danger on that score. Consider, however, the question of deletion. Forward deletion produces (98), with the Orphan Accusative. But what prevents backward deletion, which would produce *(102)? The ungrammaticality of *(102) leads to the following conclusion:

(106) The No-Backward-Deletion Constraint

Deletion does not operate backward (from right to left) in co-ordinate structures.

(106) is needed in the grammar of Slovenian to prevent *(102). But this constraint is exactly what Hankamer (1971) postulates is a linguistic universal, on totally independent grounds. The distribution of the Orphan Accusative in co-ordinate structures thus furnishes additional support for this proposed universal principle.

12. Identity of Sense Pronominalization in Universal Grammar

In many languages, Identity of Sense Pronominalization appears to be a simple process of deletion. In Spanish, for example, we find paradigms like:

(107) Ramón tiene dos caballos negros, y Manuel tiene
has two horses black and has
uno blanco.
one white
'Ramón has two black horses, and Manuel has a white one.'

(108) Ramón tiene dos caballos negros, y Manuel tiene uno.
'Ramón has two black horses, and Manuel has one.'

In the derivation of (107), the NP *caballo* seems to have been deleted, and in (108), caballo negro has been. In Slovenian, on the other hand, it has been shown that examples of Identity of Sense Pronominalization that appear to be simple deletion actually involve an intermediate stage of derivations with a pronoun, and that the pronoun is subsequently deleted. This led us to suggest, in §11.1, that deletions of a noun phrase under identity with another noun phrase universally pass through a pronominal stage. If this is correct, (107) and (108) in Spanish are not derived by simple deletion, but also go through a pronominal stage like that shown for Slovenian.

The claim that noun phrase deletion under identity universally goes through a pronominal stage has two testable empirical consequences. First, if the claim is correct there will be no languages with direct evidence against a pronominal stage in noun phrase deletion under identity. Second, there should be some languages other than Slovenian that provide evidence for an intermediate pronominal stage in Identity of Sense Pronominalization. While Spanish furnishes no evidence either for or against a pronominal stage in Identity of Sense Pronominalization, as far as we are aware, one language that does provide evidence for this pronominal stage is Czech.[16]

In Czech, Identity of Sense Pronominalization paradigms are much like those in Slovenian (cf. (30)), with the pronoun remaining in surface structure if it has no modifiers. In the presence of modifiers, the pronoun is generally deleted, as in

(109) Tomás má černý kabát, a Honza má bílý.
has black coat and has white
'Tomás has a black coat, and Honza has a white one.'

In informal, colloquial style, however, it is possible to say:

(110) Tomás má černý kabát, a Honza ho má bílý.
has black coat and it has white
'Tomás has a black coat, and Honza has a white one.'

Similarly, we find sentence pairs like:

(111) a. Tonda má košili s proužkama, zatímco Ferda ji
has shirt with stripes whereas it
má jednobarevnou.
has solid-colored

b. Tonda má košili s proužkama, zatímco Ferda má
jednobarevnou.
'Tonda has a shirt with stripes, whereas Ferda has a solid-colored one.'

In (110) and (111a), the pronouns ho and ji have been left in the surface string, whereas in (109) and (111b) they have been deleted. The conditions under which deletion takes place do not concern us here. The point that is of interest is that whereas in Slovenian the pronouns that we postulated as playing a role in Identity of Sense Pronominalization show up only with the modifier ves 'all', in Czech they appear in surface structure in a wider class of environments. Both languages furnish evidence for a pronominal stage in Identity of Sense anaphora.

French also provides interesting evidence in this connection. Whereas Spanish exhibits what appears to be deletion in (107) and (108), in French we find:

(112) Jean-Pierre a deux chevaux noirs, et Maurice
has two horses black and
en a un blanc.
of-them has a white
'Jean-Pierre has two black horses, and Maurice has a white one.'

(113) Jean-Pierre a deux chevaux noirs, et Maurice
has two horses black and
en a un.
of-them has one
'Jean-Pierre has two black horses, and Maurice has one.'

That is, the French sentences are just like those in Spanish, except that in addition French has the clitic pronoun *en* in preverbal position. This *en* is a pro-PP of the form *de*+NP.[17] This suggests that underlying *un cheval* 'a horse' is *un de chevaux* 'one of the horses'.[18] What is of interest here, however, is the fact that although the pronominal form that shows up in Identity of Sense Pronominalization in French is a pro-PP rather than a pro-NP, as in Slovenian and Czech, the stage of derivations involving this pro-form, which is only an intermediate stage in most cases in Slovenian and Czech, is the one that appears on the surface in French.

As long as linguists are content to describe each language completely in its own terms, a multiplicity of different descriptions will result. Identity of Sense anaphora in Spanish will be treated as purely a deletion phenomenon, while in French, Czech, and Slovenian pronominalization will be involved. As a result, a linguistic theory built on such descriptions will allow too much latitude for individual languages to differ, and as a result the theory will be too weak, doing little to constrain the notion 'human language' and thus making few substantive claims about the nature of language. When a linguist is dealing with a particular phenomenon in one language, he must therefore try to see how this phenomenon manifests itself in other languages and try to arrive at a canonical form for the process that will be universal. This universal statement should leave as little room for languages to differ as is consistent with the known facts, thereby constraining most strictly the notion 'human language'. This is what we are attempting to do here by positing that Identity of Sense anaphora universally works by means of Pronominalization (whatever the exact nature of this phenomenon may be) and possible subsequent deletion of pro-forms, as has shown to be necessary in any event for Slovenian (and Czech). It remains to see whether this proposal can handle the full range of facts involving Identity of Sense anaphora in human languages, including phenomena like the pro-form one(s) in English and Identity of Sense anaphora involving constituents other than noun phrases.

13. Relative Pronouns

Studies of relativization in generative grammar have generally treated relative pronouns, such as which in English, as arising from underlying full noun phrases by attachment of WH and Pronominalization, or, equivalently, by addition of the features [+WH] and [+Pro] to full noun phrases.[19] Low-level spellout rules have been assumed to give these pronominal WH constituents phonological shape as which, who, etc.

In some languages, there is morphological evidence that the relative pronoun is not a replacement for an entire noun phrase, but is merely a modifier. In these languages, the relative pronoun has adjective endings rather than noun endings. This is the case in Slovenian, as the following examples show:

(114) a. človek, brez katerega ne bi dobila dovoljenja

'the person without whom she would not have received permission'

b. človek, kljub kateremu je končno dobila dovoljenje

'the person in spite of whom she finally received permission'

c. človek, s katerim je govorila dolgo časa

'the man with whom she spoke for a long time'

The relative pronoun *kateri* has the adjectival endings -*ega*, -*emu*, and -*im* in (114) for the genitive, dative, and instrumental cases, respectively;[20] the corresponding endings for nouns would be -*a*, -*u*, and -*om*. The relative pronoun *kateri* must therefore be a modifier of the relativized noun, and not a replacement for that noun itself.

Further evidence for this conclusion is provided by the Orphan Accusative. This can be seen in the following examples:

(115) a. princip, v katerega verjamem

'the principle in which I believe'

b. *princip, v kateri verjamem

(116) a. okraj, v katerega sem prišel

'the district to which I came'

b. *okraj, v kateri sem prišel

The relative pronoun *kateri* must be in the Orphan Accusative form, as in (115a) and (116a), and not in the ordinary accusative, as in

*(115b) and *(116b).[21] Since the Orphan Accusative arises through Concord with a pronominal head, the relative pronoun *kateri* must be a modifier that agrees with the head, and not the head itself. Relativization must therefore involve not Pronominalization of an underlying noun phrase in such a way that it ends up as the relative pronoun *kateri*, but rather deletion of an underlying noun phrase in such a way that the modifier *kateri* is left behind. The fact that *kateri* appears in the Orphan Accusative, as in (115-116), shows that the deletion of the relativized noun phrase, like the other deletions discussed in this paper, passes through a pronominal stage.

That the *interrogative* pronoun *kateri* 'which' is a modifier is more obvious. It modifies the head in surface structure, having adjective endings that agree with the head in case, gender, and number.

(117) Kateri ječmen hočete?
'Which barley do you want?'

The conclusion that the relative pronoun *kateri* is also a modifier gives it the same status as the interrogative pronoun *kateri*. Surely this is not an accident. The fact that interrogative modifiers such as *which* in many languages also serve as 'relative pronouns' shows that there is a cross-linguistic generalization to be captured here.

It is possible, of course, that in some languages Relativization deletes the relativized noun phrase, passing through a pronominal stage, and in others it converts it into a relative pronoun. But in the absence of any facts that force us to such a weak theory, there is no reason to adopt it. We therefore propose that it is a universal of human language that, if a language has a relativization rule, the relativized noun phrase is pronominalized. This pronominalized noun phrase may then subsequently undergo deletion. Since there are languages in which the pronominalized noun phrase appears in surface structure, universal grammar must make the subsequent deletion of the pronominalized noun phrase optional. Particular languages may or may not have it, and it may be optional or obligatory in particular languages, or its application may be governed by other, language-particular conditions. What we are proposing, then, is that Relativization can involve pronominalization and, depending on language-particular circumstances, subsequent deletion of the relativized noun phrase, and that this pronominalization and deletion can leave behind a modifier like which that has traditionally been called a 'relative pronoun'.

14. Internal Structure of the Noun Phrase

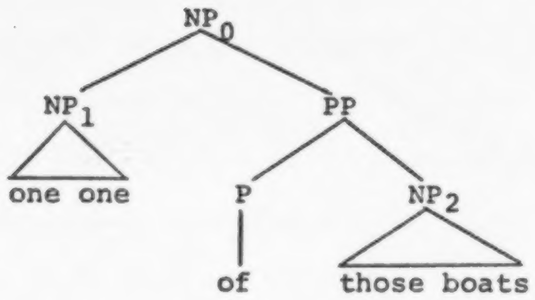
14.0 Introduction

In §14.1 it is shown how the Orphan Accusative provides evidence that the structure underlying noun phrases such as

(119) one of those boats

is roughly:

(120)



in which the second instance of one in NP$_1$ is the prop-word one that shows up in Identity of Sense Pronominalization in English. The important point is that noun phrases such as (119) are derived from structures that contain two noun phrases, as in (120).[22]

In §14.2 through §14.6, further evidence from several languages is presented to show that underlying structures like (120) are able to account for a number of facts about noun phrases that would otherwise be mysterious. Of course, one could always devise an ad hoc way of accounting for each fact in isolation, but what is interesting is that by postulating that structures like (120) universally underlie noun phrases, it is possible to account for the whole range of ostensibly disparate facts by means of a single hypothesis.

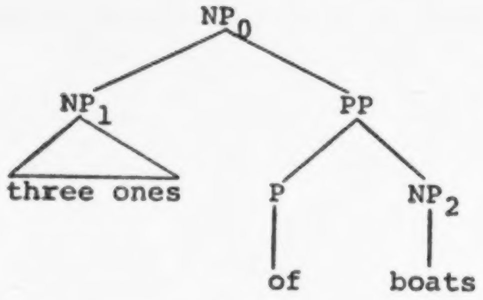
So that the full range of facts presented here can be accounted for, structures like (120) are postulated to underlie not only noun phrases with surface structures like (119), but also those like

(121) a. a boat

b. three boats

The structure underlying (121b) would be roughly:

(122)



We propose that there is a process of NP Reduction that reduces (122) to the derived structure (121b). Although the exact mechanism of this process does not concern us here, it seems likely that it consists of substitution of NP$_2$ for the pro-form one(s) in (122).

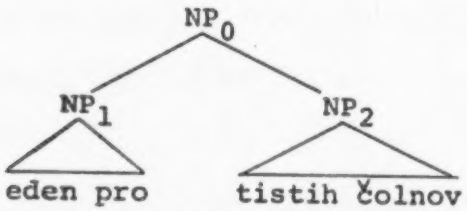
In Slovenian, the equivalent of (119) is

(123) eden tistih čolnov

'one of those boats'

in which, instead of a prepositional phrase, we find the noun phrase tistih čolnov in the genitive case. If this genitive form is taken to be basic, the Slovenian equivalent of (122) will be:[23]

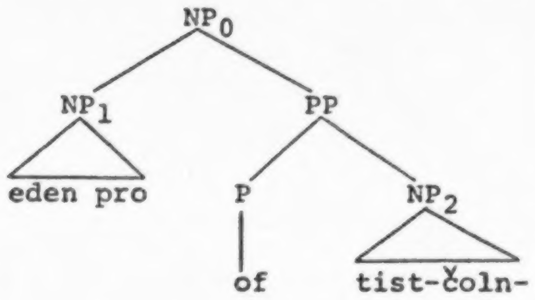
(124)



But (124) does not make NP$_2$ subordinate, as it is in (120); in (124), NP$_1$ and NP$_2$ are co-ordinate. To avoid this undesirable consequence, we propose that the cases other than nominative and accusative in languages with case be treated in the same way that prepositional and postpositional phrases are treated in languages with such overt phrases, and will represent the genitive in

Slovenian phrases like (123) as a prepositional phrase:

(125)



This presupposes a rule of Case Formation in Slovenian that will convert the PP into a genitive NP, deriving (123) from (125).

The structure that is like (125) except that it lacks the demonstrative tist- 'that' will undergo NP Reduction, producing

(126) en čoln
'a boat'

just as (121b) was derived from (122) in English.

We will not be concerned here with the exact mechanism of these rules or of derivations that include them, but rather with the kind of facts that show structures like (120), (122), and (125), which contain more than one noun phrase, to underlie surface noun phrases like (119), (121), (123), and (126).[24]

14.1 The Orphan Accusative in Slovenian

The Orphan Accusative furnishes evidence for underlying structures like (125) because of sentences like

(127) Videl sem enega tistih čolnov.
seen one those boats
'I saw one of those boats.'

in which enega 'one' is in the Orphan Accusative. Use of an ordinary accusative would be ungrammatical:

(128) *Videl sem eden tistih čolnov.

Since the Orphan Accusative arises through Concord with the pronominal head *ga*, *enega tistih čolnov* in (127) must pass through a stage in its derivation in which it is *enega ga tistih čolnov*. This means that *ga* is the head of a noun phrase with the modifier *enega*, just as *čolnov* is the head of a noun phrase with the modifier *tistih*. The NP *enega ga tistih čolnov* must therefore have an internal structure with two noun phrases, as is provided by the underlying structure (125).

14.2 *Reduction of ONE to AN in English*

In Perlmutter (1970) it is argued that the indefinite article in English is a reduction of the numeral *one* in proclitic position in the noun phrase. Stress on *one* prevents it from reducing to *an*, as in

(129) one boat

But the fact that reduction is possible in this position, as in

(130) a boat

shows that proclisis is possible here. The notion 'proclitic' must be defined in linguistic theory. We will not attempt a definition here, but will assume only that an essential element of the correct definition will involve the statement that a constituent cannot be proclitic if there is no following constituent for it to lean on. Thus, if the rest of the noun phrase is deleted, leaving only the numeral *one*, *one* has nothing to lean on and therefore cannot be proclitic. It therefore cannot reduce to *an*. As a result we find:

(131) a. Mike has three horses and Jerry has one.

b. *Mike has three horses and Jerry has an.

The final NP in (131) is left with the structure

(132) $[_{NP}$ one $]_{NP}$

and therefore cannot reduce.

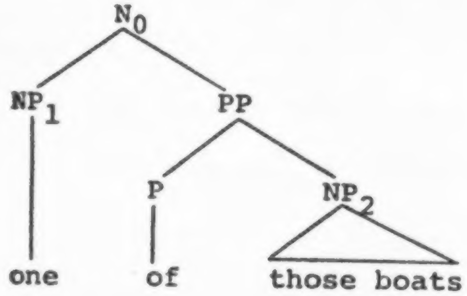
Now note the paradigm:

(133) a. one of those boats

b. *an of those boats

Why can't one reduce to an in (133)? The structure (120) provides an explanation. With (120) underlying (133), deletion of the pro-form one in NP_1 results in the derived constituent structure

(134)



The numeral one constitutes an entire NP, and therefore is not proclitic. As a result, it cannot reduce in (133) for the same reason that it cannot reduce in (131). The underlying structure (120) thus provides an explanation of this otherwise mysterious fact.

14.3 Reduction of EDEN to EN in Slovenian

The Slovenian morpheme en corresponds both to the numeral one and to the indefinite article of English, being able to appear both stressed and stressless.

(135) en čoln

'one boat, a boat'

In the masculine singular nominative, though, there are two forms of the numeral one: en and eden. If the numeral is alone in its noun phrase, we find eden instead of en:

(136) a. En čoln je prišel pravočasno, in eden ni.
one boat arrived on time and one didn't
'One boat arrived on time and one didn't.'

b. *En čoln je prišel pravočasno, in en ni.

The distribution of eden and en in Slovenian is somewhat different from that of one and an (the indefinite article) in English.

(137)

	English	Slovenian
Standing alone	one	eden
Modifying, under stress	one	en
Modifying, stressless	an	en

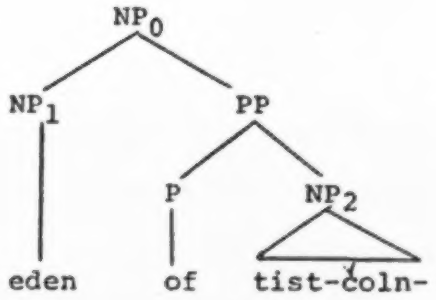
The fact that eden appears when it is standing alone in the noun phrase provides evidence for two underlying noun phrases in noun phrases like one of those boats. Note the paradigm:

(138) a. Eden tistih čolnov je prišel pravočasno.
'One of those boats arrived on time.'

b. *En tistih čolnov je prišel pravočasno.

With (125) as the structure underlying such noun phrases, deletion of the pro-form in NP_1 leaves the derived structure

(139)



in which *eden* by itself constitutes an entire noun phrase. We therefore find *eden* instead of *en* in (138) for the same reason that we find it in (136). The structure (125) provides an explanation of this fact.

14.4 *Numeral Classifiers*

Many languages have *classifiers* that must accompany numerals in the noun phrase. These classifiers are typically pro-forms for a semantic class of nouns. Thus, there may be a classifier for stick-shaped objects, a classifier for flat objects, and so on. The number and semantic range of classifiers differ from language to language; what is constant is their appearance together with numerals. What is needed in linguistic theory is a non ad hoc way of accounting for these numeral classifiers. The hypothesis that noun phrases have underlying structures like (120) provides just this. So far, we have discussed only examples in which the head of NP$_1$ was a general pro-form like *one* in English. But if a wider range of head nouns can appear in NP$_1$, we automatically have the numeral classifiers. We propose, then, that numeral classifiers are simply the heads of NP$_1$ in structures like (120). In fact, such forms exist in English as well, as in the quantified expressions

(140) a. an ear of corn

b. twenty head of cattle

c. three rashers of bacon

Here *ear*, *head*, and *rasher* perform the same function as do numeral classifiers in the languages that have them. The dependency of choice of noun on physical shape can also be seen in English in examples like the following:

(141) a. a stick of chewing gum

b. a stick of dynamite

(142) a. a sheet of paper

b. a sheet of steel

The universal structure (120) for noun phrases automatically provides the structure both for these expressions in English and for classifiers in other languages.

14.5 *EN in French*

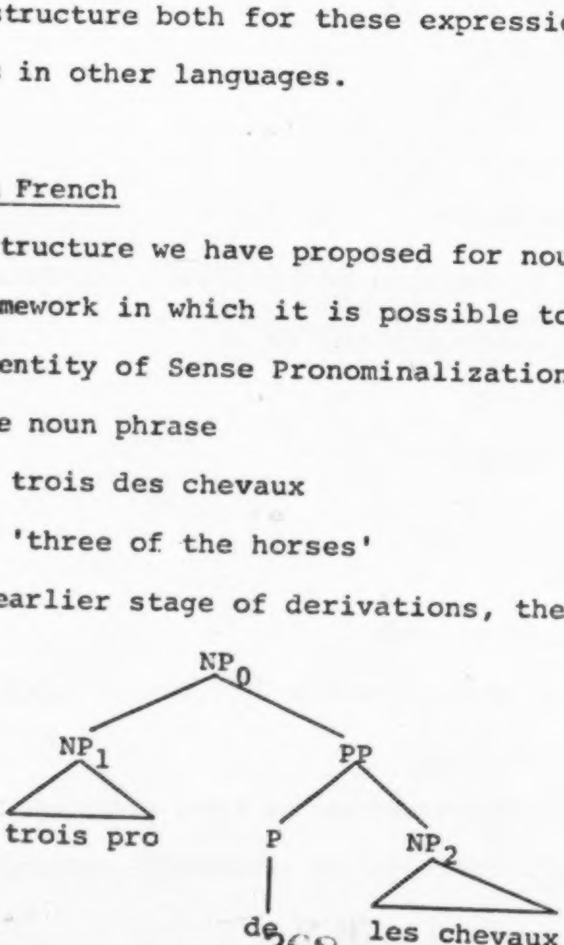
The structure we have proposed for noun phrases also provides a framework in which it is possible to explain why *en* appears in Identity of Sense Pronominalization in French, as noted in §12. The noun phrase

(143) trois des chevaux

'three of the horses'

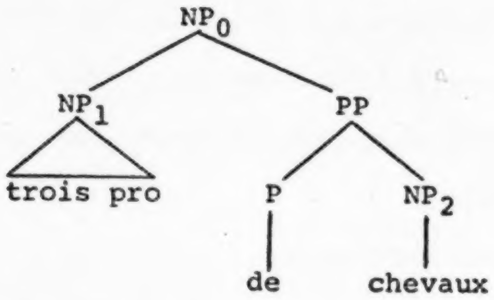
has, at an earlier stage of derivations, the structure

(144)



Note that the noun phrase in the prepositional phrase is definite: *les chevaux* 'the horses'.[25] There must therefore be another structure in which this NP is indefinite:

(145)



Whereas deletion of the pro-form in NP_1 reduces (144) to (143), the same operation, if applied to (145), would produce an ungrammatical string:

(146) *trois de chevaux
'three of horses'

Instead, we find:

(147) trois chevaux
'three horses'

(147) must therefore be derived from *(146). Similarly, the structure that would otherwise end up as

(148) *un de chevaux
'one of horses',

is reduced to

(149) un cheval
'one horse; a horse'

The rule of NP Reduction proposed in §14.0 will accomplish this reduction.[26]

With the underlying structure we have proposed for *trois chevaux* and *un cheval*, there is an automatic explanation of why en appears in Identity of Sense Pronominalization in French. Underlying the sentence

(150) Jean-Pierre a un cheval, et Maruice en a trois.
'Jean-Pierre has a horse, and Maurice has three.'

is the structure (145) in the second conjunct. Now, what is particular to French is the fact that it has the pronoun en which replaces structures of the form de+NP and is moved to pre-verbal position with the other clitics. But given the existence of this pronoun in French, the proposed universal structure for noun phrases automatically provides a source for en in (150). The ostensibly strange appearance of en in Identity of Sense Pronominalization in French is thus completely regular.

14.6 Pronominal Anaphora

The structure that we have proposed as underlying noun phrases also provides a framework in which certain initially puzzling facts of pronominal anaphora can be accounted for. We will restrict ourselves here to examples from English, but analogous facts can also be found in other languages.

In the sentence

(151) Jerry has three horses, and Roger has two of them.

the pronoun them refers to horses. Note, however, that them is not coreferential with the antecedent horses, but is rather understood in a generic sense; Roger has two horses, but not two of the same horses Jerry has. Now consider the sentence

(152) Jerry has a horse and Roger has two of them.

Here, too, the pronoun them refers to horses, in the generic sense. Neither (151) nor (152) contains, at least in surface structure, a generic antecedent *horses* for *them* to refer to, and in (152) there is not even a plural antecedent to which *them* can refer. But the plural pronouns *they* and *them*, when anaphoric, require plural antecedents. Thus, in

(153) a. The girl ran in, but I didn't see them.

b. The girl ran in, and they started to scream.

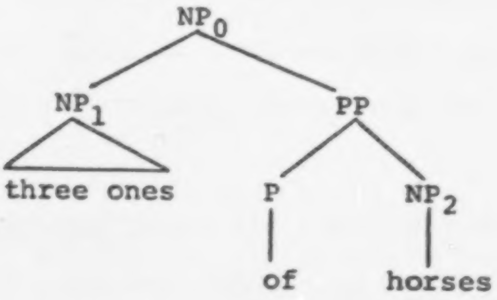
them and *they* are not anaphoric, since there is no plural antecedent for them to refer to. A feminine singular pronoun in the same position could be anaphoric, since it could refer to the feminine singular noun phrase *the girl*:

(154) a. The girl ran in, but I didn't see her.

b. The girl ran in, and she started to scream.

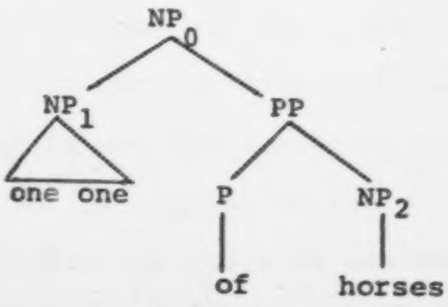
How is it, then, that *them* in (151) and (152) are anaphoric, and understood as generic? The structure that we have proposed for noun phrases can account for this phenomenon. Underlying *three horses* in (151) is a structure like

(155)



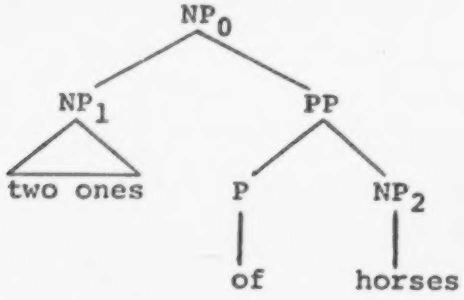
and underlying *a horse* in (152) a structure like

(156)



The underlying noun phrase in the second conjunct of both (151) and (152) is something like

(157)



The generic *horses* in the prepositional phrase in (155) and (156) provides an antecedent for the plural pronoun *them* in (151) and (152).

Note that this phenomenon is not restricted to phrases like *two of them*. In the sentence

(158) Jerry has a horse because they don't cause much pollution.

the plural pronoun *they* again refers to a generic noun phrase *horses*, which antecedent is provided by the underlying structure (156). The interesting question concerning the conditions under which NP$_2$ structures can serve as a pronominal antecedent and the conditions under which they cannot we leave for future research.

14.7 *Summary*

It has not been our purpose here to go into detail on the rules that convert underlying structures like (120), (122), and (125) into grammatical surface structures. What is important here is the fact that the occurrence of the Orphan Accusative in examples like (127) in Slovenian led us to hypothesize that what appear to be simple noun phrases in surface structure actually contain two constituent noun phrases at a deeper level of representation. Hypothesizing that such structures exist universally, we were able to explain a number of ostensibly disparate facts in several languages.

Conclusion

Language-Particular Rules and Explanation in Syntax

It has been the purpose of this paper to investigate a language-particular phenomenon--the Orphan Accusative in Slovenian. We sought to determine how this phenomenon should be accounted for in the grammar of Slovenian, and to separate what was particular to Slovenian in the Orphan Accusative from what was more general. We then considered some of the implications of our analysis for general linguistic theory.

The most important point of this paper has been to show that when one is confronted with a language-particular fact in syntax, although the course of least resistance would be to write a language-particular rule to state it, such language-particular rules give rise to questions like Questions One-Ten which a proper account of the phenomenon must answer. Unless such questions are answered, the phenomenon in question has not been explained. In the case of the Orphan Accusative in Slovenian, it has been shown that under the correct analysis, the only thing that is particular to Slovenian is the marking [+Animate] on pronouns. Everything else concerning the Orphan Accusative follows from more general principles that automatically answer Questions One-Ten, thereby providing an explanation of the facts that gave rise to them.

More difficult is the problem of how to construct linguistic theory so that incorrect solutions that raise questions like Questions One-Ten are excluded in principle. In the case of the Orphan Accusative, we were faced with a choice between the Orphan Accusative Rule (21) and marking pronouns as [+Animate]. One might argue that an explicit simplicity metric would choose the latter solution over the former, since the former involves adding an additional rule to the grammar. But at the present stage of the study of syntax, any talk of a simplicity metric is quite premature, since there are very few concrete proposals concerning the relative 'cost' of different available grammatical devices. In the case at hand, there is no explicit proposal for a simplicity metric according to which a grammar with animacy markings on pronouns is 'simpler' than one with no such markings and an additional rule. One proposal that bears on the issues raised in this paper is the Condition Principle proposed in the Epilogue to Perlmutter (1971). Conditions on transformations raise questions like Questions One-Ten, and the Condition Principle dictates the choice of a solution that involves no conditions on transformations over a solution with such conditions. In this paper, however, we have shown that questions of this kind can arise not only from conditions on rules, but also from the statement of a rule itself. This shows that the Condition Principle by itself, or even its stronger form, the No-Condition Principle, is too weak to rule out in principle solutions that give rise to questions like Questions One-Ten.

What seems to be the most promising way of approaching this problem appears in some recent work by Postal (1970a), Ross (1970), Bach (1971) and Hankamer (1971). These studies take the view that languages are not free to construct transformations at will (confining themselves, to be sure, to certain universally prescribed formal devices), but rather that there is a fixed inventory of transformations available to the grammars of particular languages, which they may or may not actually make use of. Transformations therefore will not vary from one language to another. The effect of this proposal is to go beyond the attempt to constrain transformations by purely formal means, and to place strong substantive constraints on transformations as well. In this theoretical framework, the question of the Orphan Accusative Rule (21) would not even arise, for language-particular rules of this kind would be ruled out on general theoretical grounds. The fact that it would rule (21) out on principled grounds is a merit of the theory.

In conclusion, we hope to have shown that any language-particular transformation whose structural description is rich in stating where it does and does not apply is suspect. Such rules are suspect precisely because in stating under what circumstances they do or do not apply, they fail to explain why it is that they apply where they do and not elsewhere. In the absence of a general theory that predicts what rules can and cannot state, what a rule states it cannot explain.

Footnotes

16. We are indebted to Karel Kovanda for this information about Czech.

17. Evidence for this is provided by Kayne (1969).

18. This proposal will be modified somewhat in §14.

19. This is essentially the proposal made by Smith (1961). It has been assumed in generative grammar ever since, as can be seen in Ross (1967), for example, the only change being that since the appearance of Chomsky (1965), Smith's way of doing relativization has been conceived of in terms of the syntactic features [+WH] and [+Pro] being added to the relativized noun phrase. Of relevance to the point being made here is the analysis of Kuroda (1969), in which the relative pronouns who and which are related to the interrogatives who and which.

20. The preposition brez 'without' takes its object in the genitive case, kljub 'in spite of' governs the dative, and z 'with' (written s before a voiceless consonant) governs the instrumental.

21. The preposition v in these examples takes its object in the accusative case.

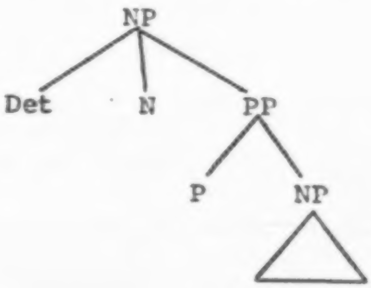
22. While we claim that (119) is derived from (120), we do not claim that (120) is necessarily the 'deepest' level of representation; (120) could itself be a derived rather than basic structure.

23. Certain English formatives, such as pro and of, are used in

these underlying structures. They are intended as symbols for entities that must be defined in linguistic theory for the use of grammars of particular languages.

24. Jackendoff (1968) deals with structures underlying noun phrases like those in (119). Although he does not say so explicitly, it appears that he intends his proposal to cover only noun phrases with overt quantifiers, although he gives no reason why noun phrases without overt quantifiers in surface structure should not have the same underlying structure. His proposal resembles that proposed here for all noun phrases in certain respects. He proposes underlying structures of the form:

(i)



which are similar to (120) and (122). The chief defect of this proposal is that Jackendoff's *Det* and *N* nodes do not together constitute an NP. In a phrase such as *a box of candy*, then, *a box* would not be an NP--a counterintuitive result that fails to capture the sameness of the phrase *a box* in *a box of candy* and other instances of *a box*. Jackendoff's structure would therefore make it impossible to capture the generalizations captured by the structures proposed here: the generalization between the operation of Concord to produce the Orphan Accusative in *enega tistih čolnov* and the operation of Concord to produce the Orphan Accusative in NPs (see §14.1), or to explain the failure of *one* to reduce in *one of those boats* (see §14.2) and of *eden* to reduce in *eden tistih čolnov* (see §14.3) by virtue of *one* and *eden* being alone in their respective NPs.

25. *de* _ + *les* becomes *des*.

26. Although we do not wish to go into the details of the mechanism of NP Reduction, it should be noted that simply deleting the pro-form in NP_1 and the preposition would be inadequate, since in examples in Slovenian and English in which the quantifier in NP_1 is *eden* and *one*, respectively, such deletion would leave it alone in NP_1, thereby preventing reduction of *eden* to *en* in *en čoln* and reduction of *one* to the indefinite article in *a boat*. That is why we suggested in §14.0 that NP Reduction probably consists of substituting NP_2 for the pro-form in NP_1 with concomitant deletion of the preposition. It should also be noted that if the quantifier of NP_1 is one of a small set that includes *one* and *each* (and their equivalents in various languages) the noun in head position after NP Reduction must be singular.

References

Andrews, A. (1971), "Case agreement of predicate modifiers in ancient Greek", Linguistic Inquiry 2.127-151.

Bach, E. (1971), "Questions", Linguistic Inquiry 2.153-166.

Bierwisch, M. and K. Heidolph (1970), Progress in Linguistics (Mouton and Co., The Hague).

Chomsky, N. (1965), Aspects of the Theory of Syntax (The MIT Press, Cambridge, Mass.).

Hankamer, J. (1971), Constraints on Deletion in Syntax (unpublished doctoral dissertation, Yale University).

Jackendoff, R. (1968), "Quantifiers in English", Foundations of Language 4.422-442.

Kayne, R. (1969), The Transformation Cycle in French Syntax (unpublished doctoral dissertation, MIT).

Kuroda, S.-Y. (1969). "English relativization and certain related problems", in Reibel and Schane (1969).

Perlmutter, D. (1970), "On the article in English", in Bierwisch and Heidolph (1970).

Perlmutter, D. (1971), Deep and Surface Structure Constraints in Syntax (Holt, Rinehart and Winston, New York).

Postal, P. (1966), "On so-called 'pronouns' in English", in Report of the Seventeenth Annual Round Table Meeting on Linguistics and Language Studies (Georgetown University Press, Washington). Reprinted in Reibel and Schane (1969).

Postal, P. (1970a), "The method of universal grammar", in P. Garvin (ed.), Method and Theory in Linguistics (Mouton and Co., The Hague).

Postal, P. (1970b), "On coreferential complement subject deletion", Linguistic Inquiry 1.439-500.

Reibel, D. and S. Schane (1969), Modern Studies in English (Prentice-Hall, Englewood Cliffs, New Jersey).

Ross, J.R. (1967), Constraints on Variables in Syntax (unpublished doctoral dissertation, MIT).

Ross, J.R. (1970), "Gapping and the order of constituents", in Bierwisch and Heidolph (1970).

Smith, C. (1961), "A class of complex modifers in English", Language 37.342-365.

Toporišič, J. (1970), Slovenski knjižni jesik 4 (Založba Obzorja, Maribor).

Evidence for Shadow Pronouns in French Relativization*

David M. Perlmutter

0. Introduction

This paper is essentially a report on work in progress in which I attempt to support the following claims as universals of human language:

(1) Rules that "chop" constituents over variables in the sense of Ross (1967) do not exist.

(2) Rules that appear to be "chopping rules" are actually "copying rules" that leave behind a shadow pronoun in the position of the constituent that has apparently been "chopped".

(3) Shadow pronouns undergo whatever transformations would normally apply to pronouns in the structures in which they appear.

(4) Shadow pronouns are subsequently deleted by a rule I will refer to as Shadow Pronoun Deletion or simply as Shadow Deletion.

(5) It is the rule of Shadow Deletion that is sensitive to island constraints in the sense of Ross (1967).

These claims, taken together, constitute a theory of movement over variables and island constraints that makes specific empirical predictions. In this paper, I will illustrate how this theory predicts a number of facts about relativization in French. I will limit myself here to examples in which the relativized noun phrase is the object of a preposition.

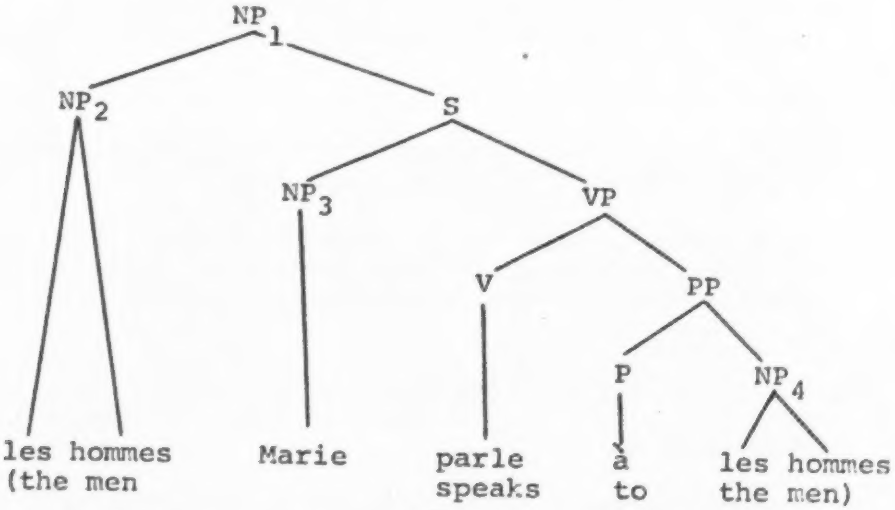
It is possible to distinguish two different theories under which relativization is accomplished by a chopping rule in the sense of Ross (1967). The two theories differ with respect to whether the relativized noun phrase is represented in underlying structure as a full noun phrase or as a pronoun. The former theory, which is that assumed by Ross (1967), would assign to a phrase like

(6) les hommes à qui Marie parle

'the men to whom Marie is speaking'

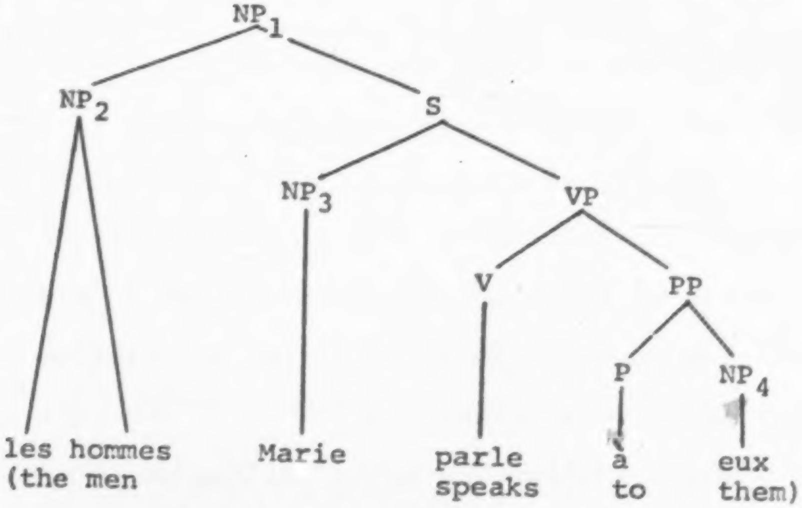
an underlying structure like[1]

(7)


NP1
NP2
S
NP3
VP
V
PP
P
NP4
les hommes
(the men
Marie
parle
speaks
à
to
les hommes
the men)

Under the latter theory, its underlying structure would be roughly

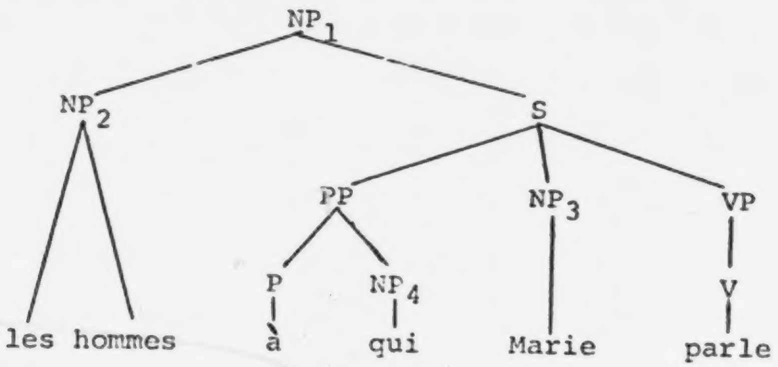
(8)


NP1
NP2
S
NP3
VP
V
PP
P
NP4
les hommes
(the men
Marie
parle
speaks
à
to
eux
them)

In either theory, the rule of Relative Movement would convert NP_4

into a relative pronoun and move it (with the preposition à) to clause-initial position to produce the derived structure

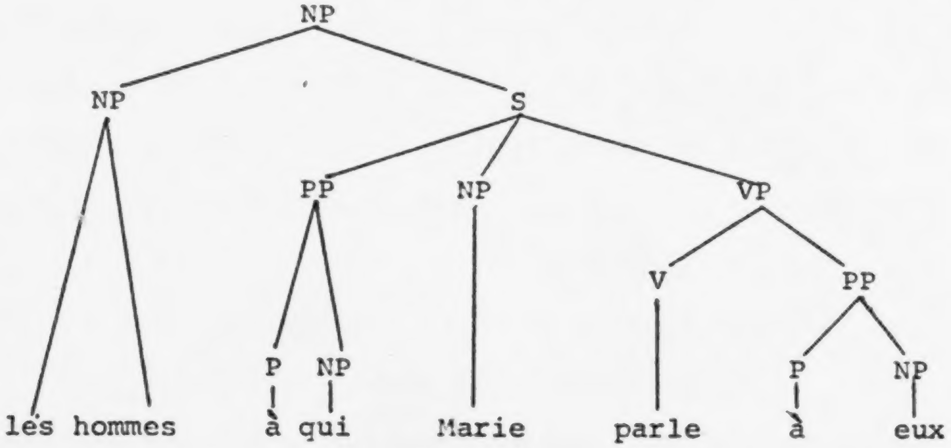
(9)


NP1
NP2
S
PP
NP3
VP
P
NP4
V
les hommes
à
qui
Marie
parle

Under both theories, NP_4 is 'chopped', that is, moved away without a pronominal copy being left behind. The theory which derives (9) from (7) in this way I will refer to as the Full NP Chopping Hypothesis, and that which derives it from (8) I will call the Pronoun Chopping Hypothesis.

I argue here that Relative Movement is not a chopping rule but a copying rule that leaves behind a pronominal copy of the moved constituent. Thus, regardless of whether (9) is derived from (7) or from (8), at some stage of the derivation the structure will look like

(10)


NP
NP
S
PP
NP
VP
V
PP
P
NP
P
NP
les hommes
à qui
Marie
parle
à
eux

with a pronominal copy of the preposed constituent left behind. Such pronominal copies I will refer to as shadow pronouns. Being pronouns, they will undergo whatever rules normally apply to pronouns. This theory of relativization I will refer to as the Shadow Pronoun Hypothesis.

The evidence for shadow pronouns in French relativization that I present here is based on the interaction of relativization and floating quantifiers. Following Paul Postal, I will refer to the rule that moves quantifiers off their noun phrases as Quantifier Float, or simply as Q-Float.

1. Floating Quantifiers in Relative Clauses and Cliticization

In French, relative clauses such as

(11) ces femmes, à qui j'ai parlé

'these women, to whom I spoke'

in which the object of a preposition has been relativized, can contain floating quantifiers:[2]

(12) a. ces femmes, à qui j'ai parlé à toutes

b. ces femmes, à qui j'ai à toutes parlé[3]

'these women, to all of whom I spoke'

In (12), à toutes 'to all' has been moved off the relativized noun phrase that ultimately ends up as the preposed constituent à qui ('to whom'). But this is not possible with all prepositions. Thus, the grammaticality of

(13) ces femmes, avec qui j'ai parlé

'these women, with whom I spoke'

contrasts with the ungrammaticality of the corresponding relative clause with a floating quantifier:

(14) a. *ces femmes, avec qui j'ai parlé avec toutes

b. *ces femmes, avec qui j'ai parlé toutes

c. *ces femmes, avec qui j'ai avec toutes parlé

d. *ces femmes, avec qui j'ai toutes parlé

'these women, with all of whom I spoke'

It is the same with devant ('in front of'). Compare

(15) ces femmes, devant qui j'ai parlé

'these women, in front of whom I spoke'

with

(16) a. *ces femmes, devant qui j'ai parlé devant toutes

b. *ces femmes, devant qui j'ai parlé toutes

c. *ces femmes, devant qui j'ai devant toutes parlé

d. *ces femmes, devant qui j'ai toutes parlé

'these women, in front of all of whom I spoke'

Why are floating quantifiers possible in a relative clause if the relativized constituent is preceded by the preposition à, but not if it is preceded by avec or devant (or any other preposition)?

Examination of these relative clauses as they would be if they were sentences in isolation provides no clue, for with a full noun phrase in place of the relativized constituent none of them allow floating quantifiers. That is, while

(17) J'ai parlé à ces femmes.

'I spoke to these women.'

is grammatical, it is impossible to have a floating quantifier in this sentence:

(18) a. *J'ai parlé à ces femmes à toutes.

b. *J'ai à toutes parlé à ces femmes.

It is the same with avec and devant. Corresponding to

(19) J'ai parlé avec ces femmes.

'I spoke with these women.'

and

(20) J'ai parlé devant ces femmes.

'I spoke in front of these women.'

there are no grammatical sentences with floating quantifiers:

(21) a. *J'ai parlé avec ces femmes (avec) toutes.

b. *J'ai (avec) toutes parlé avec ces femmes.

(22) a. *J'ai parlé devant ces femmes (devant) toutes.

b. *J'ai (devant) toutes parlé devant ces femmes.

If we examine the sentences that are just like these except that a pronoun has been substituted for the full noun phrase object of the preposition, however, the picture changes. With avec and devant, floating quantifiers are still impossible. That is, although

(23) J'ai parlé avec elles hier soir.

'I spoke with them last night.'

and

(24) J'ai parlé devant elles hier soir.

'I spoke in front of them last night.'

are grammatical, the corresponding sentences with floating quantifiers are ungrammatical.

(25) a. *J'ai parlé avec elles hier soir (avec) toutes.

b. *J'ai (avec) toutes parlé avec elles hier soir.

(26) a. *J'ai parlé devant elles hier soir (devant) toutes.

b. *J'ai (devant) toutes parlé devant elles hier soir.

With à, however, the situation is different. The sentence that is like (17) except for the fact that a pronoun has been substituted for ces femmes is not

(27) *J'ai parle à elles.
'I spoke to them.'

but rather

(28) Je leur ai parlé.
'I spoke to them.'

That is, the prepositional phrase à elles undergoes the rule of Cliticization, becoming the clitic pronoun leur that is placed in front of the verb.[4] And (28) allows floating quantifiers:[5]

(29) a. Je leur ai parlé à toutes.
b. Je leur ai à toutes parlé.
'I spoke to them all.'

We see here that two phenomena are correlated: Cliticization of pronominal objects and Quantifier Float. This suggests that floating quantifiers are possible in (29) because the prepositional object undergoes Cliticization, while they are not possible in *(25) and *(26) because the would-be source of the floating quantifiers remains as the object of a preposition in surface structure. Although I cannot go into details of the mechanism of Quantifier Float and its ordering with respect to Cliticization here, the generalization can be stated roughly as follows:[6]

(30) Quantifiers cannot float off a noun phrase that does not undergo Cliticization.

The bulk of this paper is devoted to providing further empirical support for (30) and to showing that if facts about floating quantifiers in relative clauses are to be accounted for in a non ad hoc manner, the relativized noun phrase must undergo Cliticization.

Each argument has roughly the following form:

(31) If the relativized noun phrase does not undergo Cliticization, the grammar must include an ad hoc statement to state a fact about floating quantifiers in relative clauses. But if the relativized noun phrase undergoes Cliticization, the fact in question follows automatically.

In §2, I show that under neither the Full NP Chopping Hypothesis nor the Pronoun Chopping Hypothesis is it possible for the relativized noun phrase to undergo Cliticization, and that these theories of relativization must therefore be abandoned. Under the Shadow Pronoun Hypothesis, however, the shadow pronoun left behind by Relative Movement undergoes Cliticization, and as a result the facts about floating quantifiers in relative clauses follow automatically. In §3-5, I present further evidence that the relativized noun phrase undergoes Cliticization. For the reasons given in §2, this evidence makes it necessary to adopt the Shadow Pronoun Hypothesis over the two chopping hypotheses.

2. Chopping versus Copying Theories of Relative Movement

2.1 Inadequacy of the Full NP Chopping Hypothesis

Under the Full NP Chopping Hypothesis, the relativized noun phrase is a full noun phrase that is "chopped", leaving no pronominal copy. As a result, it is at no stage of the derivation a pronoun that can undergo Cliticization. Under the Full NP Chopping Hypothesis, then, it is impossible to make the contrast between (12), on the one hand, and *(14) and *(16), on the other, follow from that between (29) and *(25) and *(26). That is, it is necessary to add an ad hoc constraint to the grammar along the following lines:

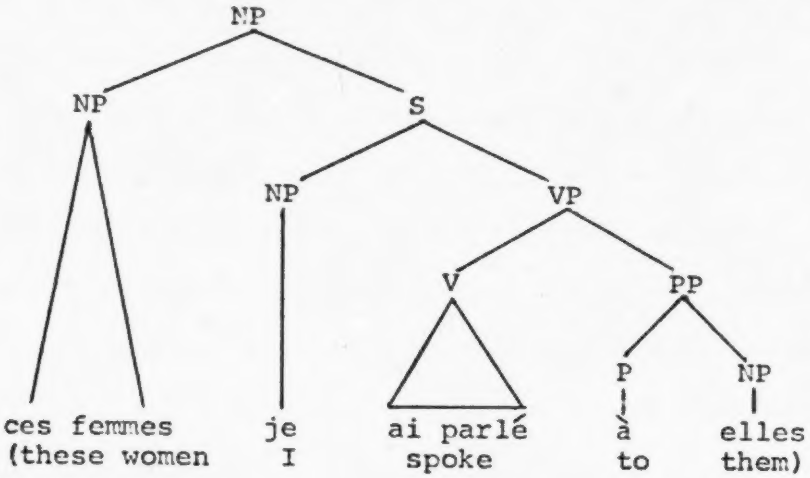
(32) Quantifiers can float off relativized objects of à, but not off relativized objects of avec or devant.

Because it is unable to use the generalization (30) to predict the occurrence of floating quantifiers in relative clauses, the Full NP Chopping Hypothesis must be abandoned.

2.2 Inadequacy of the Pronoun Chopping Hypothesis

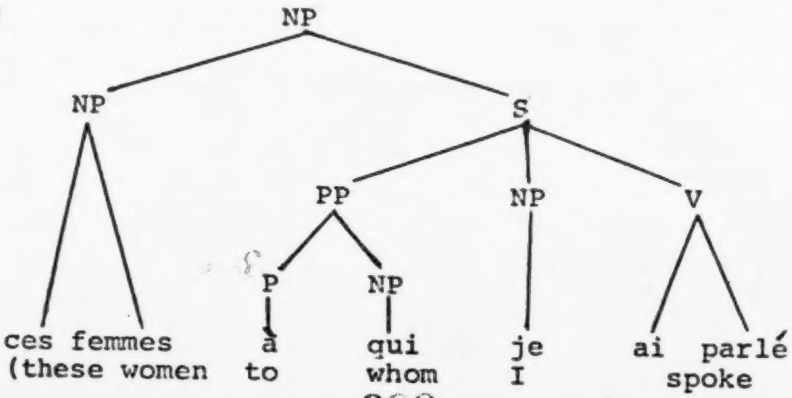
Consider now the implications of these facts for the Pronoun Chopping Hypothesis. Under this hypothesis, the relativized noun phrase is a pronoun in underlying structure, and is converted into a relative pronoun and moved to clause-initial position, leaving no pronominal copy behind. Thus, from the structure

(33)


NP
NP
S
NP
VP
V
PP
P
NP
ces femmes
(these women
je
I
ai parlé
spoke
à
to
elles
them)

the relative clause

(34)


NP
NP
S
PP
NP
V
P
NP
ces femmes
(these women
à
to
qui
whom
je
I
ai parlé
spoke

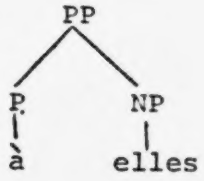
is derived. The relativized noun phrase is represented as a pronoun (*elles*) in underlying structure and as a relative pronoun (*qui*) in surface structure.

The crucial flaw in this hypothesis as well is the fact that at no point in the derivation can the relativized noun phrase undergo Cliticization. There are two possibilities: Relative Movement must take place either before Cliticization or else after Cliticization.

If Relative Movement takes place before Cliticization, the relativized constituent is already represented as *qui* in the prepositional phrase *à qui* when Cliticization applies, and it consequently cannot undergo Cliticization, which does not apply to relative pronouns. Note further that, if one attempted to change the statement of Cliticization in some ad hoc way so that it would apply to relative pronouns, it would then be necessary to invent another ad hoc operation to decliticize the cliticized relative pronoun, since it ends up not as a clitic, but in the prepositional phrase *à qui* in surface structure. It is necessary to conclude that Relative Movement does not take place before Cliticization.

If, on the other hand, Relative Movement does not take place until after Cliticization, it will be impossible to derive the phrase *à qui* from the already cliticized *leur*. Cliticization in these cases takes as input structures of the form

(35)



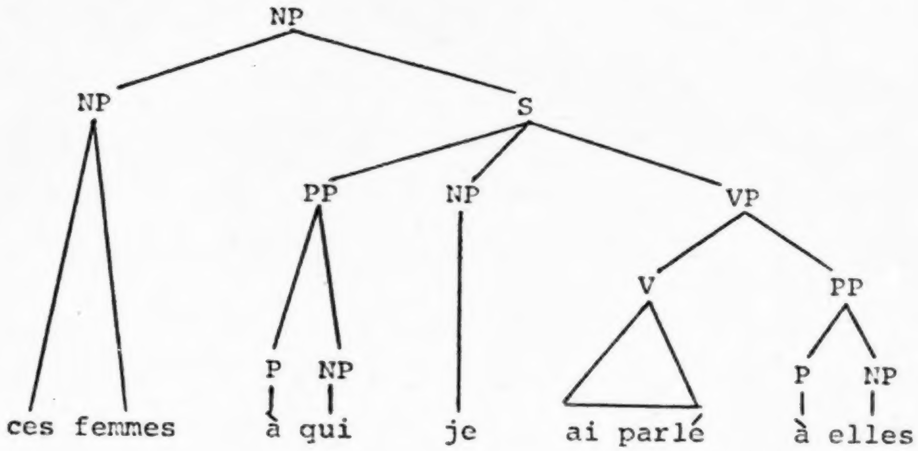
and yields as output clitic pronouns such as leur, which do not have the internal structure of prepositional phrases. In other words, Cliticization destroys the PP-structure of (35). In order to derive surface structures such as (34), in which à qui has the structure of a prepositional phrase, Relative Movement must take as input something that has PP-structure; it cannot create that structure from a structureless clitic. Thus, Relative Movement must take as input pre-Cliticization structures such as (33), and therefore cannot take place after Cliticization.

It has been shown, then, that under the Pronoun Chopping Hypothesis, it is impossible for Cliticization to apply to the relativized noun phrase, for Relative Movement can neither precede nor follow Cliticization in derivations.[7] Yet, if the grammar is to avoid an ad hoc statement like (32), the same facts must be made to follow from the generalization (30), and this can only be done if the relativized noun phrase undergoes Cliticization. The Pronoun Chopping Hypothesis must therefore be abandoned.

2.3 Adequacy of the Shadow Pronoun Hypothesis

If Relative Movement is not a chopping rule but a copying rule, no such problems arise. Taking as input either (33) or a structure in which the relativized noun phrase is represented as a full noun phrase (under this hypothesis it makes no difference), it will yield as output the structure[8]

(36)



in which the relativized noun phrase has left behind a shadow pronoun (elles). This shadow pronoun will undergo Cliticization in the normal way, producing the structure[9]

(37) [ces femmes, à qui je leur ai parlé]

Given this structure, the possibility of floating quantifiers in (12) follows automatically. In the corresponding derivation in which the relativized noun phrase has a universal quantifier, at this stage of the derivation Q-Float can apply, yielding

(38) a. [ces femmes, à qui je leur ai parlé à toutes]

b. [ces femmes, à qui je leur ai à toutes parlé]

Then Shadow Pronoun Deletion will delete the shadow pronoun, producing the grammatical surface forms

(39) ces femmes, à qui j'ai parlé

'these women, to whom I spoke'

in derivations like (37) in which Q-Float has not applied, and

(40) a. ces femmes, à qui j'ai parlé à toutes

b. ces femmes, à qui j'ai à toutes parlé

'these women, to all of whom I spoke'

in derivations like (38) in which it has.

It is the purpose of this paper to provide evidence in support of this copying theory of Relative Movement in French, and thereby to support the more general theory that no language has "chopping rules" that move noun phrases across variables, and that apparent chopping rules are actually copying rules that are followed by a rule of Shadow Pronoun Deletion. The strategy of argumentation that I follow here can be outlined as follows:

(41) a. Find a process that only pronouns can undergo. In this case it is Cliticization.

b. Find another process that depends on the first one. In this case it is Quantifier Float, which is possible only if the noun phrase off which quantifiers float undergoes Cliticization.

c. Show the facts of Q-Float in relative clauses can be accounted for in a non ad hoc manner only if Relative Movement is a copying rule, leaving behind a shadow pronoun that undergoes Cliticization and thereby makes Q-Float possible in the relative clause.

It is thus crucial to my argumentation that Q-Float indeed depend on Cliticization. Much of the paper is therefore devoted to establishing the validity of this, which has been expressed as the generalization (30). In §3-4 I provide additional evidence that Q-Float in relative clauses depends on Cliticization of the shadow pronouns left behind by Relative Movement, and in §5 it is shown that the position of floating quantifiers in relative clauses depends on the position of the shadow pronoun after Cliticization. These sections therefore all furnish evidence that Relative Movement in French is not a chopping rule but a rule that leaves behind a pronominal copy of the moved constituent.

3. Further Evidence that Q-Float Depends on Cliticization

In this section it is shown that although Cliticization is a necessary condition for Q-Float, it is not a sufficient condition, and that generalizations will be missed in a grammar of French in which Relative Movement is a chopping rule rather than a copying rule.

3.1 à versus de

Although Q-Float can apply if an object of à is relativized, it cannot apply if the relativized noun phrase is the object of the preposition de. Thus, while

(42) ces femmes, de qui je suis fier
'these women, of whom I am proud'

is grammatical, Q-Float is not possible here:

(43) a. *ces femmes, de qui je suis fier de toutes
b. *ces femmes, de qui je suis fier toutes
c. *ces femmes, de qui je suis de toutes fier
d. *ces femmes, de qui je suis toutes fier
'these women, of all of whom I am proud'

And there is a clitic, en, that corresponds to de + pronoun.[10]

(44) J'en suis fier.
'I am proud of them.'

If Relative Movement were a chopping rule, it would be necessary for the grammar to contain a constraint something like

(45) If an object of à is relativized, quantifiers can float off the relativized noun phrase, but if an object of de is relativized they cannot.

But under the Shadow Pronoun Hypothesis, floating quantifiers in relative clauses float off the pronominal copy of the relativized constituent. This hypothesis therefore correctly predicts the ungrammaticality of *(43), for Q-Float is not possible from en.

(46) a. *J'en suis fier de toutes.

b. *J'en suis fier toutes.

c. *J'en suis de toutes fier.

d. *J'en suis toutes fier.

'I am proud of them all.'

With shadow pronouns in the derivation of relative clauses, the mechanism that generates (44) will also generate (42). Since quantifiers cannot float in *(46), they will be unable to float in the derivation of *(43) either. The grammar need say nothing specifically about floating quantifiers in relative clauses. The fact that the only prepositional objects that allow quantifiers to float away in relative clauses are objects of à follows from the facts of sentences with clitic pronouns.

3.2 Two Kinds of à

Not all objects of à allow quantifiers to float away in relative clauses. Consider the following pair of sentences:

(47) J'ai parlé à ces femmes.

'I spoke to these women.'

(48) J'ai pensé à ces femmes.

'I thought about these women.'

Both form relative clauses

(49) ces femmes, à qui j'ai parlé

'these women, to whom I spoke'

(50) ces femmes, à qui j'ai pensé

'these women, about whom I thought'

but only the former allows floating quantifiers in the relative clause.

(51) a. ces femmes, à qui j'ai parlé à toutes

b. ces femmes, à qui j'ai à toutes parlé

'these women, to all of whom I spoke'

(52) a. *ces femmes, à qui j'ai pensé à toutes

*ces femmes, à qui j'ai à toutes pensé

'these women, about all of whom I thought'

The ungrammaticality of *(52) is due to the fact that pronominal complements of *penser*, while they are of the form *à + pronoun*, cliticize to *y* rather than to *leur*.[11] Thus we get

(53) J'y ai pensé.

'I thought of them.'

rather than

(54) *Je leur ai pensé.

And quantifiers cannot float off *y*:

(55) a. *J'y ai pensé à toutes.

b. *J'y ai à toutes pensé.

'I thought about them all.'

If Relative Movement were a chopping rule, the grammar would have to contain an additional constraint along the following lines:

(56) Quantifiers can float off relativized objects of *parler à* but not off relativized objects of *penser à*.

But this constraint would duplicate information that must be in the grammar anyway, to the effect that the former cliticize to *leur* and its singular counterpart *lui* while the latter cliticize to *y*, and that quantifiers can float off the dative clitic *leur*, but not off *y*.[12] With shadow pronouns in the derivation of relative clauses, the pronominal copy of the relativized noun phrase actually undergoes Cliticization. As a result, it becomes *leur* in the derivation of (51) but *y* in the derivation of *(52). The fact that quantifiers can float off the relativized noun phrase in (51) but not in *(52) follows automatically from the fact that they can float off *leur* but not off *y*.

4. Q-Float from Adjectival Complements and its Interaction with Subject Raising into Object Position

Sentences such as:

(57) Les hommes sont fidèles à ces femmes.

'The men are faithful to these women.'

form relative clauses

(58) ces femmes, à qui les hommes sont fidèles

'these women, to whom the men are faithful'

and Q-Float is possible in these relative clauses.

(59) a. ces femmes, à qui les hommes sont fidèles à toutes

b. ces femmes, à qui les hommes sont à toutes fidèles

'these women, to all of whom the men are faithful'

Sentences like (57) can also be embedded beneath verbs such as *croire* 'I believe' that allow Subject Raising into object position.[13]

(60) J'ai cru les hommes fidèles à ces femmes.

'I believed the men (to be) faithful to these women.'

(60) can also be relativized:

(61) ces femmes, à qui j'ai cru les hommes fidèles

'these women, to whom I believed the men (to be) faithful'

But in (61), Q-Float is impossible:

(62) a. *ces femmes, à qui j'ai cru les hommes fidèles à toutes

b. *ces femmes, à qui j'ai à toutes cru les hommes fidèles

'these women, to all of whom I believed, the men (to be) faithful'

If Relative Movement were a chopping rule, it would be necessary to include in the grammar a constraint something like:

(63) Quantifiers cannot float off a relativized dative complement of an adjective whose subject undergoes Subject Raising into object position.

But with shadow pronouns in the derivation of relative clauses, this fact follows automatically from another. Q-Float is possible in (59) because a pronominal complement of fideles in (57) can cliticize

(64) Les hommes leur sont fidèles.

'The men are faithful to them.'

and Q-Float off the cliticized complement is possible.

(65) a. Les hommes leur sont fidèles à toutes.

b. Les hommes leur sont à toutes fidèles.

'The men are faithful to them all.'

Since (65) represents a stage in the derivation of (59), (59) is grammatical as well.

If the subject is raised into object position, however, Cliticization of the adjectival complement is impossible.

(66) a. *J'ai cru les hommes leur fidèles.

b. *Je leur ai cru les hommes fidèles.

'I believed the men (to be) faithful to them.'

As a result, Q-Float is impossible. Under the Shadow Pronoun

Hypothesis, the derivation of *(62) is possible only if, at an earlier stage of derivations, the complement of _fidèles_ is cliticized to _leur_, making Q-Float possible. Under this hypothesis, then, the ungrammaticality of *(62) follows automatically from the impossibility of Cliticization in *(66).[14]

5. The Position of Floating Quantifiers in Relative Clauses

French has a rule, which Kayne (1969b) calls L-TOUS, that moves floating quantifiers to the left of their source under certain conditions. This rule applies with some matrix verbs but not with others. Thus, L-TOUS applies with _vouloir_ 'want', for example, so that from the structure underlying

(67) Martin a voulu leur parler à tous.
'Martin wanted to speak to them all.'

the sentence

(68) Martin a à tous voulu leur parler.
'Martin wanted to speak to them all.'

is also derived. In (68), _à tous_ has been moved to the left of its source (_leur_). But if the matrix verb is _croire_ 'believe', on the other hand, L-TOUS does not apply. Thus, while

(69) Martin a cru leur avoir parlé à tous.
'Martin believed he had spoken to them all.'

is grammatical,

(70) *Martin a à tous cru leur avoir parlé.

in which _à tous_ has been moved to the left of its source (_leur_), is not.[15] As the Shadow Pronoun Hypothesis predicts, the positions that floating quantifiers can occupy in relative clauses are exactly the same as those that are possible in sentences with clitic pronouns. Thus, while both

(71) ces soldats, à qui Martin a voulu parler à tous

'these soldiers, to all of whom Martin wanted to speak'

and

(72) ces soldats, à qui Martin a à tous voulu parler

'these soldiers, to all of whom Martin wanted to speak'

are grammatical, in relative clauses with croire as the matrix verb, only the first possibility is grammatical.

(73) ces soldats, à qui Martin a cru avoir parlé à tous

'these soldiers, to all of whom Martin believed he had spoken'

(74) *ces soldats, à qui Martin a à tous cru avoir parlé

If we exclude from consideration sentences to which L-TOUS applies, the following generalization holds of the remaining sentences of French:

(75) A floating quantifier must be to the right of its source (the noun phrase off which it has floated).

The generalization (75) makes possible another empirical test of the Shadow Pronoun Hypothesis. This hypothesis states that shadow pronouns are present in pre-surface structures of relative clauses, and that they undergo Cliticization and therefore will occupy the positions that Clitics normally occupy before they are deleted by Shadow Pronoun Deletion. The Shadow Pronoun Hypothesis and the generalization (75) together therefore predict that if clitics move to different positions in different sentences in which L-TOUS does not apply, this will be reflected in the possible positions of floating quantifiers in relative clauses. And there are sentences of French that make it possible to test this prediction. The sentences

(76) J'ai laissé Paul écrire des lettres à ces enfants.
'I let Paul write letters to these children.'

and

(77) J'ai laissé écrire des lettres à ces enfants par Paul.
'I let Paul write letters to these children.'

differ in their transformational history; in the derivation of (77), the rule that Kayne (1969b) calls FAIRE-Attraction has applied while in that of (76) it has not. Both sentences can be relative clauses:

(78) ces enfants, à qui j'ai laissé Paul écrire des lettres
'these children, to whom I let Paul write letters'

(79) ces enfants, à qui j'ai laissé écrire des lettres par Paul
'these children, to whom I let Paul write letters'

And both allow floating quantifiers in their relative clauses:

(80) ces enfants, à qui j'ai laissé Paul écrire des lettres à tous
'these children, to all of whom I let Paul write letters'

(81) ces enfants, à qui j'ai laissé écrire des lettres à tous par Paul
'these children, to all of whom I let Paul write letters'

But putting the floating quantifier between ai and laissé affects their grammaticality differently.

(82) *ces enfants, à qui j'ai à tous laissé Paul écrire des lettres

(83) ces enfants, à qui j'ai à tous laissé écrire des lettres par Paul
'these children, to all of whom I let Paul write letters'

If Relative Movement were a chopping rule, it would be necessary to state this fact by means of an ad hoc constraint along the following lines:

(84) A floating quantifier from the indirect object of a verb embedded beneath a laisser-type verb can move to the position after the auxiliary verb in the matrix sentence if FAIRE-Attraction has applied, but not if it hasn't.

Under the Shadow Pronoun Hypothesis, however, this follows automatically. One consequence of the application of FAIRE-Attraction is that in sentences in which it has applied, clitic pronouns from the embedded sentence move up to the finite verb of the matrix sentence. If the object à ces enfants in (76) and (77) is replaced by a pronoun, we find

(85) J'ai laissé Paul leur écrire des lettres.

'I let Paul write them letters.'

corresponding to (76), but the sentence that corresponds to (77) is

(86) Je leur ai laissé écrire des lettres par Paul.

'I let Paul write them letters.'

In (86), but not in (85), the clitic from the complement (leur) has moved up into the matrix sentence. As a result of the generalization (75), the two sentences differ in grammaticality if a floating quantifier is put between ai and laissé.

(87) *J'ai à tous laissé Paul leur écrire des lettres.

(88) Je leur ai à tous laissé écrire des lettres par Paul.

'I let Paul write letters to them all.'

*(87) is ungrammatical because the floating quantifier precedes its source, while (88) is grammatical because the floating quantifier follows its source, as it must. The same thing will account for the contrast between *(82) and (83) if Relative Movement leaves behind shadow pronouns that subsequently undergo clitic movement.

Under the Shadow Pronoun Hypothesis, no special statement is needed to account for the contrast between *(82) and (83), for it follows automatically from that between *(87) and (88).

6. Predictions of the Shadow Pronoun Hypothesis

The argument presented here has been based on the fact that Q-Float from prepositional objects in French is possible if the object is the clitic pronoun leur and impossible otherwise. Q-Float thus provides a means of showing that Relative Movement in French is not a chopping rule; a pronoun must be left behind in order to serve as the source of the floating quantifiers that appear in relative clauses.

The hypothesis that shadow pronouns appear in the derivation of French relative clauses predicts that for any other phenomenon with respect to which pronouns and full noun phrases behave differently, the behavior characteristic of pronouns will be found in relative clauses.

The claim that no language has rules that "chop" constituents over a variable without leaving a copy thus entails that for any grammatical phenomenon with respect to which pronouns behave differently from full noun phrases, interaction of this phenomenon with what appear to be "chopping rules" will exhibit the behavior characteristic of pronouns. This is a testable claim. In order to test it for a given apparent "chopping rule" in any language, it is only necessary to find some phenomenon with respect to which pronouns and full noun phrases behave differently. If sentences in which the apparent "chopping rule" applies exhibit the behavior characteristic of pronouns, the Shadow Pronoun Hypothesis must be modified or abandoned.

I will illustrate how this claim can be tested with an example from Portuguese. In Portuguese, infinitives in certain environments are inflected to agree with their subjects. Thus, we find sentence pairs like the following:

(89) a. *Vi os cavalos correr.
b. Vi os cavalos correrem.
'I saw the horses run.'

*(89a) has the ordinary infinitive correr, but in (89b) the infinitive has the plural ending -em, in agreement with its subject (os cavalos). In the corresponding relative clause, however, the situation is reversed:

(90) a. os cavalos que vi correr
b. *os cavalos que vi correrem
'the horses that I saw run'

If Relative Movement were a chopping rule, this would be a mysterious fact, and it would be necessary to add some ad hoc constraint to the grammar in order to state it. The Shadow Pronoun Hypothesis, however, correctly predicts that (90a) will be grammatical and *(90b) ungrammatical. In the corresponding sentence that is like (89) except that a pronoun has been substituted for os cavalos, we find exactly the same thing:

(91) a. Vi-os correr.
b. *Vi-os correrem.
'I saw them run.'

The infinitive does not agree with the clitic pronoun os. Under the Shadow Pronoun Hypothesis, then, the lack of infinitival agreement in (90) is no mystery; it follows automatically from the lack of agreement in (91).

In Portuguese, as in French, shadow pronouns undergo Cliticization in the normal way. The lack of agreement in (91) is a result of Cliticization, for infinitives agree with uncliticized pronouns. Thus, a pronoun modified by somente 'only' does not cliticize, and the infinitive agrees with it:

(92) a. *Vi somente êles correr.

b. Vi somente êles correrem.

'I saw only them run.'

It is only cliticized pronouns that do not trigger infinitival agreement.

These facts from Portuguese are of interest because they make it possible to extend in two ways the results on French that have been presented here.

First, the data from French has been confined to examples with non-restrictive relative clauses because restrictive relatives in which the relativized noun phrase has a universal quantifier are unnatural, as was noted in footnote 2. But examples with restrictive relatives such as (90) in Portuguese show clearly that restrictive relatives work exactly the same way.

Second, it is not possible to extend the data on floating quantifiers in French to questions, since the questioned constituent cannot have a universal quantifier. This is not restricted to French, as these English examples show:[16]

(93) a. *Who did you speak to all of?

b. *To all of whom did you speak?

(94) a. *Which women did you speak to all of?

b. *To all of which women did you speak?

But the interaction of infinitival agreement and Question Movement in Portuguese makes it possible to show that Question Movement also leaves behind a pronominal copy of the moved constituent.

(95) a. Que cavalos você viu correr?

b. *Que cavalos você viu correrem?

'What horses did you see run?'

(96) a. Quais cavalos você viu correr?

b. *Quais cavalos você viu correrem?

'Which horses did you see run?'

The fact that the infinitive in (95)-(96) is not inflected follows automatically under the Shadow Pronoun Hypothesis, under which these sentences are derived from structures with a shadow pronoun:

(97) [$\left\{\begin{matrix}\text{Que}\\ \text{Quais}\end{matrix}\right\}$ cavalos você vi-os correr?]

The clitic pronoun os in (97) does not trigger infinitival agreement, just as it does not in (91), and the lack of agreement in (95) and (96) follows automatically.

The interaction of relativization and Question Movement with infinitival agreement in Portuguese thus gives some substance to the claim that shadow pronouns are left behind not only by Relative Movement in French, but by any rule in any language that moves constituents over a variable.

One aspect of this claim requires further clarification, however. I have shown here that in French and Portuguese, the copy left behind by movement rules is a pronoun. This shadow pronoun satisfies general constraints on pronominalization in French and

Portuguese; the antecedent both precedes and commands the pronoun. I claim that apparent "chopping rules" are really copying rules. However, it may not be the case that in every language the copy is necessarily a pronoun. In a language in which pronominalization constraints are different, it is possible that the copy left behind is a full noun phrase, and the moved constituent, or the head of the relative clause, is a pronoun. I do not know whether or not this is ever the case, and leave it as an open question for future research. If it is, the algorithm for testing the claim that chopping rules are really copying rules sketched above will have to be modified accordingly. Given the pronominalization constraints of French and Portuguese, the hypothesis that chopping rules are really copying rules and the hypothesis that the copy left behind is a shadow pronoun make the same empirical predictions. In languages with different constraints on pronominalization, the two may make different empirical predictions. Such languages would then provide the crucial evidence to decide which of these two hypotheses is to be stated in linguistic theory as universal.

7. Island Constraints as Constraints on Shadow Deletion

It has been shown here that Relative Movement in French leaves behind a shadow pronoun that is subsequently deleted by the rule of Shadow Deletion. Nonetheless, this process is sensitive to island constraints in the sense of Ross (1967). Thus, the relative clause

(98) *ces Norvégiens, à qui je connais la femme que a parlé
'these Norwegians, to whom I know the woman who spoke'

is ungrammatical because it violates the Complex Noun Phrase Constraint, and

(99) *ces Norvégiens, à qui que Claude a parlé est probable

'these Norwegians, to whom that Claude spoke is probable'

is ungrammatical because it violates the Sentential Subject Constraint, the corresponding example in which Extraposition has applied being grammatical:

(100) ces Norvégiens, à qui il est probable que Claude a parlé

'these Norwegians, to whom it is probable that Claude spoke'

Ross showed that copying rules do not obey island constraints. Since Relative Movement is a copying rule, it follows that it is Shadow Deletion and not Relative Movement that is sensitive to island constraints. Island constraints therefore are not constraints on movement transformations, but rather constraints on deletion transformations of a particular type.[17]

8. Shadow Deletion as the Basis of a Typological Division Among Languages

8.0 Introduction

It has been shown here that the deletion of shadow pronouns (Shadow Deletion) is a rule distinct from the copying rules that move constituents across variables. From the claim that this distinction is universal it follows that languages can differ as to whether or not they have the rule of Shadow Deletion to delete the pronouns left behind by movement rules.

From the claim that it is Shadow Deletion and not movement rules that are sensitive to island constraints it follows that whether or not a language will be sensitive to island constraints is an automatic consequence of whether or not it has Shadow Deletion.

The consequences of these two claims provide the basis for a typological division among languages. Some empirical consequences of this typological division are illustrated here for three languages--Arabic, Japanese, and Turkish.

8.1 *Arabic*

In Arabic, shadow pronouns remain in relative clauses in surface structure.[18]

(101) Hāðihi hiya s-sayyāratu l-latī r-ražulu štara*hā*.
this dem. the-car which the-man bought-it
'This is the car that the man bought it.'
'This is the car that the man bought.'

The relative clause contains the shadow pronoun *ha* (cliticized to the verb) referring to the relativized noun phrase (*as-sayyaratu*). Arabic thus does not have the rule of Shadow Deletion in relative clauses.

The claim that it is Shadow Deletion that is sensitive to island constraints therefore entails that in Arabic it will be possible to relativize into islands. And this is in fact the case.

(102) Hāðihi hiya s-sayyaratu l-latī ?aʕrifu r-ražula
this dem. the-car which I-know the-man

l-laðī štarāhā
who bought-it
'This is the car that I know the man who bought.'
'This is the car that I know the man who bought it.'

Although (102) is an example of relativization going down into a relative clause, it is not a violation of the Complex Noun Phrase Constraint because the shadow pronoun (hā) has not been deleted, but remains in surface structure.

There are, however, some initially puzzling examples in Arabic in which relativization into a relative clause is possible even though the shadow pronoun does not appear in surface structure.

(103) Hāðā huwa r-ražulu l-laðī ra?aytu s-sayyārata l-latī
this dem. the-man who I-saw the-car which

štarāhā.
bought-it
'This is the man who I saw the car that bought-it.'
'This is the man who I saw the car that bought.'
'This is the man who I saw the car that he bought it.'

(104) Ra?aytu sayyāratan ?aʕrifu ražulan sa-yaštarīhā.
I-saw car I-know man will-buy-it
'I saw a car I know a man will buy it.'
'I saw a car I know a man will buy.'
'I saw a car that I know a man who will buy it.'

If Shadow Deletion has not applied in (103) and (104), these sentences should contain the pronoun huwa 'he' in the inner-most relative clause. And if Shadow Deletion has applied, there are two problems: first, it would be strange for Shadow Deletion in Arabic relative clauses to apply to subjects and not to objects, and second, if Shadow Deletion is sensitive to island constraints, how can sentences like (103)-(104), in which a pronoun inside a complex noun phrase has been deleted, be grammatical?

The solution to both problems emerges from the realization that Arabic has a general rule, which I will call Subject Pronoun Drop, that deletes subject pronouns in general. Thus, while the sentence

(105) Huwa štarāhā.
'He bought it.'

is grammatical, it is much more natural to delete the subject, yielding

(106) ?ištarāhā.
'He bought it.'

Similarly, alongside

(107) Huwa sa-yaštarīhā.
'He will buy it.'

the more natural alternative is

(108) Sa-yaštarīhā.
'He will buy it.'

It is Shadow Deletion, not Subject Pronoun Drop, that is sensitive to island constraints.[19] If the subject pronouns in (103)-(104) are deleted not by Shadow Deletion but rather by Subject Pronoun Drop, the apparent violations of island constraints in (103)-(104) will no longer be anomalous. Evidence that the subject pronouns in

(103)-(104) are actually deleted by Subject Pronoun Drop rather than by Shadow Deletion comes from the fact that their deletion, like that of the pronouns in (105) and (107), is optional. Alongside (103), the sentence

(109) Haðā huwa r-ražulu l-laðī ra?aytu s-sayyārata l-latī
this dem. the-man who I-saw the-car which

huwa štarāhā.
he bought-it

'This is the man who I saw the car that he bought it.'

in which the shadow pronoun huwa 'he' remains in surface structure, is also grammatical. By postulating that it is Subject Pronoun Drop that deletes the shadow pronoun huwa in (103)-(104) in exactly the same way that it does in (106) and (108), the two problems raised by (103)-(104) are solved simultaneously. Only shadow pronouns that are subjects can fail to appear in the surface structure of Arabic relative clauses because only subject pronouns are deleted by Subject Pronoun Drop, and the apparent violation of island constraints in examples like (103) and (104) constitutes not a counterexample to, but further evidence for, the claim that it is Shadow Deletion that is sensitive to island constraints.

8.2 Japanese

Kuno (forthcoming) has pointed out that in Japanese there are grammatical relative clauses in which a noun phrase inside a syntactic island has been relativized; the examples I cite are taken from Kuno.

Relative clauses in Japanese precede the head noun. Japanese does not have relative pronouns; the relativized noun phrase is simply deleted. Thus, alongside a sentence like

(110) Kodomo wa inu o kawaigatte-ita.
child dog Acc. petting-was
'The child was petting a dog.'

it is possible to form a relative clause on inu 'dog'

(111) kodomo ga kawaigatte-ita inu
child Nom. petting-was dog
'the dog that the child was petting'

Similarly, alongside the sentence

(112) Kodomo ga kawaigatte-ita inu wa sinde simatta.
child Nom. petting-was dog died perf.
'The dog that the child was petting died.'

it is possible to form a relative clause on kodomo 'child':

(113) kawaigatte-ita inu ga sinde simatta kodomo
petting-was dog Nom. died perf. child
'the child that the dog that was petting died'
'the child that the dog that he was petting died'

In (113), kodomo, a constituent of a relative clause (kodomo ga kawaigatte-ita inu) has been relativized. It is the same in

(114) kite-iru yoohuku ga yogorete-iru sinsi
wearing-is suit Nom. dirty-is gentleman
'the gentleman who the suit that is wearing is dirty'
'the gentleman who the suit that he is wearing is dirty'

(115) kite-iru yoohuku ni kodomotati ga inki o
wearing-is suit Dat. children Nom. ink Acc.

kakete simatta sinsi
spill perf. gentleman

'the gentleman who the children spilled ink on the suit that is wearing'

'the gentleman who the children spilled ink on the suit that he is wearing'

in which sinsi 'gentleman', a constituent of the relative clause

(116) sinsi ga kite-iru yoohuku
gentleman Nom. wearing-is suit
'the suit that the gentleman is wearing'

has been relativized.

These apparent island violations are not limited to the Complex Noun Phrase Constraint. Also grammatical in Japanese are relative clauses such as

(117) Ueda-san ga kaita koto ga yoku sirarete-iru bun
Mr. Nom. wrote Comp. Nom. well known-is article
'the article which that Mr. Weda wrote is well-known'
'the article which that Mr. Weda wrote it is well-known'

in which bun 'article' has been relativized out of the sentential subject in

(118) Ueda-san ga sono bun o kaita koto ga
Mr. Nom. that article Acc. wrote Comp. Nom.

yoku sirarete-iru
well known-is
'That Mr. Weda wrote that article is well-known.'

apparently in violation of the Sentential Subject Constraint.

Within a theoretical framework in which Ross's island constraints are viewed as universal constraints on the process of relativization itself, these facts would show the constraints to be not universal but language-particular, and one would be forced to list those languages that obey them and those that do not. Once it is realized that island constraints are constraints on Shadow Deletion, however, the solution is obvious: it is only necessary to postulate that Japanese, like Arabic, lacks the rule of Shadow Deletion. Some evidence for this is to be found in the fact that under certain conditions shadow pronouns in Japanese relative clauses show up in surface structure. Thus, alongside (115), the relative clause

(119) kare ga kite-iru yoohuku ni kodomotati ga
he Nom. wearing-is suit Dat. children Nom.

inki o kakete simatta sinsi
ink Acc. spill perf. gentleman

'the gentleman who the children spilled ink on the suit that he is wearing'

in which the shadow pronoun kare appears in surface structure, is also grammatical. If Japanese, like Arabic, lacks the rule of Shadow Deletion, its failure to obey island constraints will follow automatically from its lack of Shadow Deletion.

The obvious difference between Japanese and Arabic is the fact that whereas in Arabic shadow pronouns regularly show up in the surface structure of relative clauses, in Japanese they show up only rarely. If it is correct that the apparent violations of island constraints in Japanese are due to its lacking the rule of Shadow Deletion, there is only one possibility here: the relativized noun phrase in Japanese must be deleted, not by Shadow Deletion, but by another deletion rule that is not sensitive to island constraints. The situation with respect to all relativized constituents in Japanese would then be the same as that with relativized subjects in Arabic which, as was shown in §8.1, are deleted by Subject Pronoun Drop. In other words, the hypothesis presented here predicts that there must be a deletion rule in Japanese that deletes relativized noun phrases as well as others. If there is evidence in Japanese for such a deletion rule, the fact that Japanese has relative clauses that apparently violate island constraints will constitute striking evidence for the claim that in languages that obey island constraints it is Shadow Deletion that is sensitive to the constraints.

Confirming evidence comes from the fact that just as Arabic has a rule of Subject Pronoun Drop that deletes subject pronouns, Japanese has a general rule of Pronoun Drop that deletes all pronouns. Thus, alongside the sentences

(120) Kodomo ga inu o kawaigatte-ita.
child Nom. dog Acc. petting-was
'The child was petting the dog.'

and

(121) Kare ga inu o kawaigatte-ita.
he
'He was petting the dog.'

we also have

(122) Inu o kawaigatte-ita.
'He was petting the dog.'

in which the subject pronoun has been deleted. Deletion of object pronouns is also possible, producing sentences like[20]

(123) Kawaigatte-ita.

'He was petting it.'

Thus, it is only necessary to assume that at some late stage of derivations the relativized noun phrase is represented as a pronoun,[21] and it will be deleted by Pronoun Drop, which is not sensitive to island constraints.[22] The relativized noun phrase in (113) will thus be deleted in exactly the same way as the subject of (122). Similarly, the relativized noun phrase in (114) and (115) will be deleted in the same way as the subject of

(124)	Yoohuku	o	kite-iru.
	suit	Acc.	wearing-is

'He is wearing a suit.'

and the relativized noun phrase in (117) will be deleted by Pronoun Drop just as the object of

(125)	Ueda-san	ga	kaita.
	Mr.	Nom.	wrote

'Mr. Weda wrote it.'

is. Since the relativized constituent is in each case deleted by Pronoun Drop rather than by Shadow Deletion, the apparent violations of island constraints in surface structure follow automatically.

It should now be clear what the universal prediction is that is made by the hypothesis that it is Shadow Deletion that is sensitive to island constraints:

(126) If shadow pronouns are deleted by Shadow Deletion, island constraints will not be violated. If they are not deleted, or are deleted by some other rule, relativization into islands will be possible.

Crucially, no language can have another rule that deletes shadow pronouns alone. Whether or not shadow pronouns can be deleted by another rule in a given language depends on whether or not the language has an independently motivated rule that will automatically delete shadow pronouns as well as others.

8.3 Turkish

Turkish has a rule of Pronoun Drop that deletes both subject and object pronouns.[23] Thus, alongside

(127) O arabayı aldı
he car bought
'He bought the car.'

we have the more natural sentence

(128) Arabayı aldı.
'He bought the car.'

with the subject pronoun deleted. Similarly, both

(129) Adam onu aldı.
man it bought
'The man bought it.'

and

(130) Adam aldı.
'The man bought it.'

with the object pronoun deleted, are grammatical. Since (127)-(130) are all grammatical, Pronoun Drop in Turkish is optional.[24]

Turkish relative clauses precede the head. The relativized noun phrase is deleted and the verb is made into a participle with the ending -En if the relativized noun phrase is the subject and -dIK if it is an object.[25] Thus, if the sentence

(131) Adam arabayı aldı.
man car bought
'The man bought the car.'

is made into a relative clause on adam 'man' we get

(132) arabayı alan adam
car buy-part man
'the man that bought the car'

with the subject participle alan, and if it is a relative clause on araba 'car' we get

(133) adamın aldığı araba
man-Gen. buy-part car
'the car that the man bought'

Crucially, shadow pronouns can not appear in Turkish relative clauses.

(134) *o arabayı alan adam
he car buy-part man

(135) *adamın onu aldığı araba
man-Gen. it buy-part car

*(134) and *(135) contrast with (127) and (129).

Since Pronoun Drop is optional, and shadow pronouns in relative clauses are deleted obligatorily, they cannot be deleted by Pronoun Drop. Under the hypothesis presented here, it follows that Turkish has the rule of Shadow Deletion in relative clauses. Since Shadow Deletion is sensitive to island constraints, the hypothesis immediately predicts that it will be impossible to relativize into syntactic islands in Turkish. And this prediction is correct. Thus, a sentence like

(136) Arabayı alan adamı seviyorum.
car-Acc. buy-part man-Acc. I-like
'I like the man who bought the car.'

cannot be a relative clause on araba 'car':

(137) *alan adamı sevdiǧim araba
buy-part man-Acc. like-part car
'the car that I like the man who bought (it)'

Under our hypothesis, the ungrammaticality of *(134)-*(135) is immediately translated into the prediction that *(137) will be ungrammatical.

In this connection it should be noted that while the hypothesis advanced here immediately translates the ungrammaticality of *(134)-*(135) into the prediction that relativization into syntactic islands will be impossible in Turkish, if this hypothesis is embedded in the current theory of transformational grammar, the ungrammaticality of *(137) does not follow automatically from that of *(134)-*(135). The reason for this is simple: in the theory of transformational grammar, rules are ordered and apply in sequence.

If the Shadow Pronoun Hypothesis is embedded in transformational theory, then, it is necessary to make an additional statement in order to make the ungrammaticality of *(137) follow from that of *(134)-*(135); it is necessary to state that in the grammar of Turkish, Shadow Deletion is ordered so as to apply before Pronoun Drop. If Pronoun Drop applied before Shadow Deletion, shadow pronouns in relative clauses would be deleted by Pronoun Drop, which is not sensitive to island constraints, as happens in Japanese. As a result, examples like *(137) would be grammatical.

It is possible to obtain the correct result within transformational theory by ordering Shadow Deletion before Pronoun Drop, or by making Shadow Deletion cyclical and Pronoun Drop post-cyclical. But to do that is to miss the point that the so-called 'ordering' of these rules is predictable; if Pronoun Drop applied first, there would be no motivation for a rule of Shadow Deletion. Thus, no language can have a grammar in which Pronoun Drop applies before Shadow Deletion; the data in any such language could be handled by a grammar without Shadow Deletion. Yet, within current transformational theory there is no way to make use of the predictability of the 'ordering' of Pronoun Drop and Shadow Deletion to avoid having to state it as an ad hoc ordering fact. The question of how the theory should be changed to remedy this defect I leave open for future research.

9. Summary

Restricting attention to examples in French in which the object of a preposition has been relativized, it has been shown here that floating quantifiers are possible in a relative clause just in case floating quantifiers are also possible in the corresponding sentence in which the relativized noun phrase has been replaced by a pronoun. If, in that sentence, floating quantifiers are impossible--because the pronoun is unable to cliticize or because the pronoun becomes a clitic off which quantifiers cannot float--floating quantifiers are impossible in the corresponding relative clause as well. In order to capture this generalization in the grammar, it is necessary that the relativized noun phrase undergo Cliticization.

In surface structure, however, the relativized noun phrase appears in the prepositional phrase à qui which has undergone Relative Movement. It cannot undergo both Cliticization and Relative Movement. If Relative Movement precedes, there is no way that its output à qui can undergo Cliticization, and if Cliticization precedes, there is no way to create the prepositional phrase à qui from the already cliticized pronoun leur. But since it is necessary that the relativized noun phrase undergo Cliticization in order to predict correctly the possibility of floating quantifiers in relative clauses, I have proposed that Relative Movement is not a chopping rule but a copying rule, and that it is the pronominal copy of the relativized noun phrase that it leaves behind that undergoes Cliticization and off which quantifiers float. I also proposed that this mechanism is universal--that all rules that appear to 'chop' constituents across variables are actually copying rules that leave behind a shadow pronoun, and that this pronominal copy is subsequently deleted by a rule of Shadow Deletion. This hypothesis predicts that in any language that has some process with respect to which pronouns and full noun phrases behave differently, the behavior characteristic of pronouns will be found in cases of apparent chopping rules. The predictions made by this hypothesis were shown to be correct for the inflected infinitive in Portuguese, which also extends the empirical support for the hypothesis to restrictive relative clauses and questions. It was then pointed out that since copying rules are not sensitive to Ross's island constraints, once chopping processes are shown to consist of a copying rule followed by a deletion rule, it must be the deletion rule that is sensitive to island constraints. It was then shown that it can be determined on independent grounds whether or not a language has the rule of Shadow Deletion in relative clauses, and that this automatically predicts whether or not relativization into syntactic islands is possible in the language in question.

References

Chomsky, N. (forthcoming) "Conditions on Transformations", to appear in S. Anderson and P. Kiparsky (eds.), *Studies Presented to Morris Halle* (Holt, Rinehart and Winston, New York).

Hankamer, J. (1971) *Constraints on Deletion in Syntax*, unpublished doctoral dissertation, Yale University.

Kayne, R. (1969a) "On the Inappropriateness of Rule Features", *Quarterly Progress Report No. 95* of the Research Laboratory of Electronics, MIT.

Kayne, R. (1969b) *The Transformational Cycle in French Syntax*, unpublished doctoral dissertation, MIT.

Kuno, S. (forthcoming) *Notes on Japanese Grammar* (The MIT Press, Cambridge, Mass.).

Neubauer, P. (1970) "On the Notion 'Chopping Rule'", in *Papers from the Sixth Regional Meeting of the Chicago Linguistics Society*, University of Chicago.

Perlmutter, D. (1971) *Deep and Surface Structure Constraints in Syntax* (Holt, Rinehart and Winston, New York).

Perlmutter, D. and J. Orešnik (forthcoming) "Language-particular Rules and Explanation in Syntax", in S. Anderson and P. Kiparsky (eds.), *Studies Presented to Morris Halle* (Holt, Rinehart and Winston, New York).

Postal, P. (forthcoming) *On Raising*.

Ross, J. (1967) *Constraints on Variables in Syntax*, unpublished doctoral dissertation, MIT.

Underhill, R. (1972) "Turkish Participles", *Linguistic Inquiry* 3.87-99.

Footnotes

* This paper owes a great deal to Paul Postal, who stimulated my interest in floating quantifiers, and to Jean-Roger Vergnaud, with whom I have had many interesting discussions about French syntax. I am indebted to Jorge Hankamer, Paul Postal, John R. Ross, and Jean-Roger Vergnaud for reading a preliminary version of this paper and making many valuable suggestions for improvement, to the native speakers of French--especially Jeannette Grillet, Alain Rouveret, and Jean-Roger Vergnaud--whose intuitions served as the source of the French data, and to the students from MIT and Harvard whose reactions to this material as it developed helped shape its development. I am also indebted to Suely Lima and Carlos Quicoli for their intuitions about Portuguese, to John Makhoul and Joan Maling for help with the Arabic examples, and to Susumu Kuno and Jorge Hankamer for the examples from Japanese and Turkish. Needless to say, none of these people necessarily agree with what is said here, and responsibility for errors is mine alone. This work was supported in part by a grant from the National Institutes of Health (Grant No. 5 TO1 HD 00111-07) and a grant from the National Institutes of Mental Health (Grant No. 5 TO1 MH 13390-05) to the Massachusetts Institute of Technology, and a grant from the United States Office of Education (OEC-0-70-4986(823)) to the Language Research Foundation.

1. All tree diagrams in this paper are immensely oversimplified, since no attention has been paid to aspects of sentence structure that are not relevant to the argument. Matters

such as the position of the verb in underlying structure and the question of whether or not structures like (7) and (8) are the 'deepest' level of representation are therefore ignored here.

2. I will use forms of tous 'all' in the examples in this paper. Chacun 'each' works the same way. My examples are appositive relatives because restrictive relatives in which the relativized noun phrase has a universal quantifier are unnatural. This is also the case in English. Compare the appositive relative

(i) these women, to all of whom I spoke

with the restrictive relatives

(ii) a. *the women to all of whom I spoke
b. *the women that I spoke to all of

3. For some speakers of French, (12b) is ungrammatical. Instead they find grammatical

(iii) ces femmes, à qui j'ai toutes parlé
'these women, to all of whom I spoke'

It appears that these speakers have an obligatory rule that deletes the preposition à when it immediately follows the auxiliary verb. Although this paper describes a variety of French which has no such deletion, its conclusions hold for both dialects. Speakers who find (12b) ungrammatical and (iii) grammatical need only remember the deletion rule in question in order to convert the dialect described here into their own.

4. For discussion of Cliticization, see Kayne (1969a, 1969b) and Perlmutter (1971, Chapter Two).

5. It should be clear throughout this paper that examples with floating quantifiers and those without them differ in underlying structure. Phraseology such as "(28) allows floating quantifiers" and "Quantifier Float is possible in (28)" are therefore to be interpreted as abbreviations for such phrases as "in the underlying structure that is like (28) except for the presence of a universal quantifier, Quantifier Float is possible."

6. Recall that this paper is restricted to Relative Movement and Quantifier Float involving objects of prepositions. (30) is not true of subject noun phrases.

7. Note that it has been shown that neither order of application of these two rules yields a satisfactory result. This means that not only is it impossible to achieve correct results by means of an extrinsic ordering statement in the grammar, but also that, even if the two rules are unordered in the grammar, or if one is cyclical and the other post-cyclical, it is impossible for the relativized constituent to undergo Cliticization if Relative Movement is a chopping rule.

8. If Relative Movement takes as input structures such as (33) in which the relativized noun phrase is a pronoun, the copy left behind by Relative Movement must necessarily be pronominal. But if structures in which the relativized noun phrase is a full noun phrase serve as input to Relative Movement, it is conceivable that the copy left behind by Relative Movement is also a full noun

phrase, and that it subsequently becomes pronominal as a result of normal processes of pronominalization. Although I refer to the 'pronominal copy' left behind by Relative Movement throughout this paper, it makes no difference to my hypothesis whether the copy is pronominal to begin with or only becomes so subsequently.

9. I use square brackets to represent underlying or intermediate stages of derivations.

10. For arguments that en is derived from structures of the form de + pronoun, see Kayne (1969b).

11. For supporting arguments, see Kayne (1969a, 1969b).

12. For some speakers of French, quantifiers can float off y, so that for these speakers *(55) are grammatical. And for these speakers, *(52) are grammatical as well, as predicted by the Shadow Pronoun Hypothesis.

13. The existence of such a rule has recently been called into question by Chomsky (1973), but Postal (forthcoming) provides overwhelming evidence for its existence. Relevant evidence from French includes the fact that the underlying subject of the complement, if pronominal, ends up as a clitic on the matrix verb, as in

(iv) Je les ai cru fidèles à ces femmes.
'I believed them (to be) faithful to these women.'

and that it undergoes Reflexivization if it is coreferential with the subject of the matrix sentence, as in

(v) Les hommes se sont crus fidèles.
'The men believed themselves (to be) faithful.'

Note further that the argument for Relative Movement being a copying rather than a chopping rule given here would go through even if there were no rule of Subject Raising into object position.

14. It is interesting to note that if the subject of the adjectival complement is raised into object position, as is the case in these examples, the sentence with a strong form pronoun is also ungrammatical:

(vi) *J'ai cru les hommes fidèles à elles.
'I believed the men (to be) faithful to them.'

Both *(66), in which Cliticization has applied, and *(vi), in which it has not, are ungrammatical. It is therefore clear that some kind of filtering device is necessary to characterize these sentences as ungrammatical. Q-Float in relative clauses sheds light on the nature of this filtering device in two ways.

First, one should ask at what level of derivations this filter applies and what it is in these sentences that causes the filter to reject them as ungrammatical. Comparison of *(66) and *(vi) with the grammatical example (61) provides evidence that bears on these questions. Since relative clauses contain a shadow pronoun prior to the application of Shadow Deletion, at some stage of the derivation (61) is represented as one of the following three structures:

(vii) a. [ces femmes, à qui j'ai cru les hommes fidèles à elles]

(vii) b. [ces femmes, à qui j'ai cru les hommes leur fidèles]

c. [ces femmes, à qui je leur ai cru les hommes fidèles]

Shadow Deletion deletes the shadow pronoun, and the resulting structure is grammatical. This shows clearly that the filter in question is sensitive to the presence of the pronoun in surface structures; if the pronoun is deleted, the result is grammatical.

Second, since it is clear that it is the pronoun that causes the ungrammaticality of *(66) and *(vi), two alternatives suggest themselves. Under the first, the pronoun undergoes Cliticization, and the filter rejects as ungrammatical a post-Cliticization structure: (vii-b) or (vii-c). Under the second alternative, Cliticization is blocked, and it is the structure (vii-a) with the strong form pronoun that is rejected by the filter. Q-Float in relative clauses provides precisely the evidence that is needed to decide between these two alternatives. If the pronoun underwent Cliticization, Q-Float would be possible in the relative clause, and examples like *(62) would be grammatical: Cliticization of the shadow pronoun would make Q-Float possible, and deletion of the offending pronoun by Shadow Deletion would enable *(62) to get by the filter. Since *(62) is ungrammatical, it is necessary to conclude that the pronoun does not undergo Cliticization, and that it is consequently structures like (vii-a) and *(vi) with strong-form pronouns that are discarded by the filter. And this is precisely what has been proposed for such cases by Kayne (1969a). Q-Float in relative clauses, together with the analysis of relative clauses presented here, thus furnishes additional evidence for a filter like that postulated by Kayne.

It should also be noted that examples like (61) show clearly that while Cliticization of the shadow pronoun is a necessary condition for Q-Float in relative clauses, it is not necessary to relativization itself. In the derivation of (61), Relative Movement preposes _à qui_, leaving behind the shadow _à elles_, which is subsequently deleted by Shadow Deletion. At no time does it undergo Cliticization, but the resulting relative clause is nonetheless well-formed.

15. One speaker has informed me that for him *(70) is grammatical. In general, there seems to be considerable variation among speakers on the positioning of floating quantifiers. However, the hypothesis tested in §5 holds for every speaker I have asked; regardless of judgments on particular sentences, the position of floating quantifiers is always the same in a relative clause as in the corresponding sentence in which the relativized constituent has been replaced by a pronoun.

16. This would follow automatically in a theory that derives question-words from forms of _some_, since _some_ and _all_ are mutually exclusive.

17. Ross's characterization of island constraints is this: "Variables in chopping rules, feature-changing rules, and unidirectional rules of deletion cannot cross island boundaries."

While the characterization 'unidirectional rules of deletion' has been called into question by Hankamer (1971), it is clear that some deletion rules obey island constraints. Under the proposal made here that chopping rules do not exist, the characterization of rules that obey island constraints becomes simpler. For a proposal under which at least some of what Ross calls 'feature-changing rules' are also deletion rules, see Perlmutter and Orešnik (forthcoming). Neubauer (1970) denies that constraints on chopping rules can be handled as constraints on deletion, but I find his evidence inconclusive. I hope to deal with it in a more comprehensive work on this subject.

18. I give first a literal translation of the sentence into English, followed by a translation in the derivation of which Shadow Deletion has applied. In cases (such as (102-104) below) in which the English translation with Shadow Deletion is unfathomable, I also give a translation in which Shadow Deletion has not applied.

19. Note that whereas Shadow Deletion is a deletion rule triggered by the moved constituent, deleting shadow pronouns across a variable, Subject Pronoun Drop does not require either another constituent to trigger the deletion or the statement of an environment in which it applies. The fact that Shadow Deletion obeys island constraints and Subject Pronoun Drop does not should follow from this difference.

20. These sentences are actually ambiguous. For example, (122) could mean 'I was petting the dog', 'They were petting the dog', etc., and the object in (123) can likewise be taken to be any pronoun.

21. This is an interesting point. Since Japanese does not have a rule of Relative Movement, shadow pronouns in Japanese can not be said to have resulted from a copying operation such as Relative Movement. The relativized noun phrase must therefore either be pronominalized by the usual pronominalization rules, if such there be, or else represented as a pronoun in underlying structure.

22. Pronoun Drop in Japanese, like Subject Pronoun Drop in Arabic, does not require either another constituent to trigger the deletion or the statement of an environment in which it applies. From this its insensitivity to island constraints should follow automatically.

23. It is not relevant to the argument whether subject and object pronouns are deleted by the same rule or two separate rules.

24. The circumstances under which it does and does not apply do not concern us here.

25. E and I represent vowels that change in accordance with rules of vowel harmony. For a discussion of these Turkish participles in a generative framework, see Underhill (1972).

Induced Errors

Nancy Stenson

1. Introduction

Research in foreign language teaching has recently seen a great proliferation of theories on the best way to improve teaching methods and increase understanding of second language acquisition. To this end, scholars have readily espoused various theoretical and/or methodological concepts as old theories and methodologies have been proved inadequate. Among the concepts which have been mentioned and discussed in the TESOL literature over the past several years, we find linguistic analysis, contrastive analysis, and error analysis. In this paper, I would like to briefly place these three types of analysis in perspective, and then treat in some detail one aspect of error analysis: those problems created by the learning situation.

1.1 Linguistic analysis, most simply stated, is the work done by linguists in providing a descriptive grammar of a language. Linguistic analysis was initially done within the prescriptive framework of Latin grammar, later during the first half of the 20th century, within the structuralist framework, and most recently within the framework of generative grammar. As stated by Chomsky (1965, page 3):

> Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance.

That is, linguistic analysis is concerned with the competence which native speakers of a language have, a competence quite different from that which a non-native learner of the language would possess. For this reason, generative linguists have expressed doubt as to the value of linguistic analysis to the language teacher. The following statement of the issue was made by Chomsky at the 1966 Northeast Conference:

> I am frankly skeptical about the significance, for the teaching of language of such insights and understanding as have been attained in linguistics and psychology. It is possible--even likely--that principles of psychology and linguistics, and research in these disciplines, may supply insights useful to the language teacher. But this must be demonstrated, and cannot be presumed. It is the language teacher himself who must validate or refute any specific proposal.

Certainly the grammar of a transformational linguist would not be useable as a teaching grammar, being far too abstract and formal to adapt to classroom use. However, the insights gained by linguistic analysis have more than once been used by pedagogues in developing materials for language teaching. Most notably the Audio-Lingual Method, still widely used in U.S. schools, was based on the tenets underlying structural linguistics. Many of the more recent insights gained by research in transformational generative grammar may likewise be useful to the language teacher, if not in providing actual teaching materials, then at least in providing a better understanding of certain of the students' problems. As research continues, insights into the learning process will hopefully enable teachers to better understand some of what goes on in the classroom from a psycholinguistic point of view.

1.2 Contrastive analysis has both a strong and a weak form (cf. Wardhaugh 1970). The strong version embodies a claim of predictive power: that a contrastive analysis of two language systems can be used to predict the errors that speakers of the first language will make in learning the second, and thus can be used directly in the development of teaching materials. As Wardhaugh points out, this position is untenable at present, for it makes demands on linguistic theory and a non-existent theory of contrastive analysis that cannot be met.[1] It is, however, the version that underlies much of the work that has been done in contrastive analysis.

The weak version approaches the problem from the opposite direction, taking the purpose of contrastive analysis to be the account of observed errors by starting with classroom data and using the differences between the two linguistic systems to explain the errors.[2] Many contrastive analysis attempts which fail in their claim to meet the demands of the strong formulation do meet at least some of the demands of the latter version--their failure is in not doing all they promise.

No theory of contrastive analysis, strong or weak, should be expected to account for all errors of language learning. Much evidence is already available which suggests that many errors are due to target-language rule deviance as well. In addition, there are many errors induced by the classroom situation, but which cannot be considered to be a function of performance.[3] There is some hope, however, that as linguists sharpen their ability to distinguish between the universals and the idiosyncrasies in a language, at least the weak version of contrastive analysis can become a more useful tool in language learning research.

1.3 Arising from the failure of contrastive analysis to adequately account for student errors, references began appearing in the literature of a new technique: error analysis. It has been proposed in several places as an alternative or supplement to contrastive analysis (cf. Buteau, Dušková, Madarasz and Banathy) without its being clearly formulated just what is meant by the term or what its proponents feel are its goals. In view of this vagueness, this issue is worthy of more careful scrutiny than it has been previously given.

Error analysis grew out of transformational linguistic theory and the notion of language as a rule-governed system. A serious and oft-noted problem of most contrastive analyses is that they typically concentrate on superficial differences between languages; more recent developments in linguistic research, on the other hand, have indicated that surface forms reveal little about the nature of language, and that more valid insights into language structure are to be found on a more abstract level, a level at which the differences and similarities discovered by contrastive analysis might prove to be irrelevant, or nonexistent.[4] Within such a framework, the error is perceived as a rule violation with respect to the target language alone. Understanding of student learning habits must depart from this point.[5]

Seen in this light, it is clear that error analysis is just that--analysis. It is NOT a methodology. Its purpose is to further our understanding of what the student knows and doesn't know about the language he is learning, of what generalizations he has made about the new language, and where he has gone wrong. Error analysis will NOT show a teacher how to correct or avoid the error.[6] It is, however, a necessary preliminary to the development of materials which can accomplish this goal.

It is also apparent when one thinks about what is involved in actually doing an error analysis, that there is a relatively small difference between it and the weak form of contrastive analysis. The contrastive/error analysis controversy had its origin in the directionality of the analysis: whether one started from the analysis of the language being studied or from the student's performance. This conflict only exists where we take contrastive analysis to mean the strong version, which claims predictive power based on description of the two languages involved.

If we compare the weak form of contrastive analysis to error analysis, we find that in fact they both make their departure from the same point: the target language as the student speaks it. Both attempt to account for observed facts. Differences arise at the next step--how to account for the data observed. Contrastive analysis looks for points of interference from the student's native language, while what has been called error analysis considers errors only in terms of the student's formulation of the target language system. These two approaches are not inconsistent but, rather, focus on different problems within the same approach. I therefore suggest that contrastive analysis (in its weak form) should be considered just one aspect of the larger area of error analysis. Moreover, both are specific forms of linguistic analysis.

In oversimplified terms, then, we can divide student errors (in both comprehension and production of the second language) into at least the following three types: errors due to incomplete acquisition of the target grammar (which may in turn be due to either inter- or intra-lingual difficulties), errors due to the exigencies of the situation (the topic to be discussed below), and errors due to normal problems of language performance.[7] The real task of error analysis from this perspective is to explain why one aspect of the target grammar has not been adequately acquired while a second is learned without difficulty. To the extent that this can be done by reliance on the linguistic systems of the two languages, this work is what we characteristically call contrastive analysis. On the other hand, there is empirical evidence that areas of contrast between two languages are not always the areas of greatest difficulty for native speakers of one learning the other. It remains for other branches of error analysis to explore and ultimately explain these other areas of difficulty. The following deals with one such aspect of language learning. Here, as in every case, the crucial first step is to find the source of the errors made by students. This is the overall task of error analysis, within which all more detailed types of analysis fall.

2. Classroom Induced Errors

In this section, I would like to characterize some types of student errors in a language classroom that result more from the classroom situation than from either the student's incomplete competence in English grammar, or first language interference. Students are easily led into making errors in the course of classroom participation by the structure of the situation; these are errors which it is doubtful they would produce in spontaneous speech. Such induced errors tell us little about the level of the student's language competence but are worthy of study in that they are easy to overlook, may easily cause inaccurate assessment of the student's ability, and, if ignored, may reinforce misunderstanding and form the basis for later problems. Any analysis of student errors must also take such mistakes into account as phenomena separate from errors of spontaneous speech.

In section 2.1, I will discuss several examples in which failure to appreciate meaning differences leads to misuse of lexical items. Section 2.2 treats unusual syntactic errors that seem to be related to the context in which they were produced rather than the grammar of the student who produced them. In section 2.3 I examine some errors that students seem trapped into by the structure of drills as well as certain traps that they (unexpectedly) escape. Section 3 treats some mistakes which occur despite every effort to structure the situation so that the correct answer should follow naturally. In sections 4 and 5 I discuss briefly some of the implications for both the teacher and linguist in the ESL classroom.

The data from which my examples are drawn was gathered in Tunis, Tunisia, during the summer of 1971 from observation of adult English classes at the Institut Bourguiba des Langues Vivantes, and from high school students during practice teaching sessions of the 1971 Peace Corps TEFL training program in Mahdia, Tunisia. Most of the students were in intermediate or advanced classes.

It is important to keep in mind during this discussion that any attempt to explain what a student has in mind when he makes an error can only be tentative at best. The unexpected nature of some of the examples presented below testifies to the highly speculative nature of such work. My purpose is not to draw fast conclusions or prescribe solutions. Rather, I wish to emphasize the importance of teachers being aware of the dangers inherent in a classroom situation where one person is in control of the material covered, and also of the difference between errors made in the structured classroom environment and those made during spontaneous speech. It is worth pointing out at the outset that underlying much of the discussion is the oft-stated plea that language teachers learn more about the language they are teaching. Not that teachers should become descriptive linguists, but rather that they should acquire early in their training and teaching experience a sound understanding of the results of linguistic studies. Many of the errors I have observed could have been avoided or easily corrected had the teacher been aware of even basic grammatical insights.

2.1 A teacher may inadvertently mislead students by the way he defines a lexical item, or by the order in which he presents material. For example, given worship as a general word for pray, the students immediately attached the same preposition to the new word that they knew to be required with the familiar one, and began speaking of "worshipping to God". Habits which may develop from such analogies between two related items seem especially difficult to break; several students made this mistake in succession, even after the teacher had corrected the first of them.[8]

The problems of both word definition and order of presentation were brought into focus when students were confronted for the first time with a single lexical item consisting of a verb + particle. They were given the definition of point out through example sentences with appropriate gestures, and then asked to use it in sentences. Those students who did not merely paraphrase the teacher's examples were all clearly treating the construction as two separate lexical items, point, which they already knew, and the preposition out. Thus, the new lexical item came out sounding to them like just another way to say point to or point at. One student with a little more imagination offered "When I see a ship in the sea, I point out", which the teacher corrected to "...I point it out to my friends." This is probably not what the student meant at all. The chances are at least equally good that he was using out as a directional: "I point out (to sea, to the ship, etc.)..." That he intended out in his sentence to indicate direction is plausible, since all students were having a great deal of trouble with point out, and since omission of a direct object pronoun where one is required was rare, if not unheard of, at least among these students. If this was the case, the teacher's insertion of the pronoun it is no help at all to the student. It doesn't help him learn the real meaning of point out, which both the teacher and he assume he knows, and if he's a very attentive student, given to quick generalizations, he might even be led, by a correction like this, to start inserting it incorrectly with other directionals. Treating verb + particle forms as a unit at some point would at least give a tool for correcting individual lexical items.

Immediately after point out, and without fully understanding it, the students were given notice and asked to use it in sentences. This led to the sentence "The barometer noticed that it wouldn't be fine." This student appears to have confused the two new vocabulary items and, since one word bears a causative relation to the other, this serves to reinforce the confusion.[9] The student might not make a mistake like this in a normal conversation--he would be more inclined to use a word he's sure of, like show, if he ever needed to talk about barometer readings--but once he has misused a word or (worse) a pair of words, even under drill conditions, he will probably, if not set straight, remain confused and make additional errors. This is not to suggest that a teacher should necessarily avoid teaching related words together. Indeed, such a device can be a very useful teaching technique, and whether to use it or not must depend on the situation and the class. However, a teacher who is attuned to such relationships between words (not always evident on the surface of things) can do much to avoid difficulties or clear them up once they present themselves.

A final type of induced vocabulary error came from students being forced to demonstrate a distinction that they were vaguely aware of, but unable to explain adequately. For example, asked to demonstrate the difference between should and must, a student said: "We should have worked in order to buy clothes, but we must have worked in order to eat." This class knew that must was in some way stronger than should, but were apparently not entirely clear on just how--in general, they steered clear of should and used must everywhere. The student who had to demonstrate the difference did indeed make a distinction of degree, but in so doing, he transferred the force of necessity from the end to the means.

A similar error arose when a student attempted to characterize the difference between at and into: "We look at the moon, but Armstrong looked into the moon." The distinction had been demonstrated solely by examples and pantomime ("Now I'm looking at the bag; now I'm looking into the bag" (moving closer to the bag to indicate peering into it)) and this student probably got the idea that the distinction was one of thoroughness or closeness rather than of surface versus interior. Just saying "no" to a student with such a misconception teaches him nothing, yet a teacher unaware of what is going on can do little else.

Situations like these can tell us a lot we wouldn't otherwise learn about what a student knows and doesn't know. It is impossible to measure his competence with structures and vocabulary he never uses. Since errors are our only way of determining what aspects of language the student has not yet acquired, one could easily be led to assume (without the help of induced errors) that non-appearance of a certain structure rather than deviant appearance indicates acquisition of that structure. In many cases nothing could be farther from the truth. The difference between systematic errors which reflect possible rule differences, interference, etc., and those errors arising from gaps in student grammar should be obvious. The latter constructions are extremely difficult to get at, since they are rarely (and never systematically) reflected in the student's speech. Thus a tremendous number of differences between a student's competence and that of a native speaker may go unobserved or unexplained because there is no data to work from. Here, then, is an area where the importance of induced errors as distinguished from spontaneous errors is apparent. The former provide otherwise unobtainable information about student competence. Two points follow from this: first, that vocabulary misuse may reflect non-acquisition of certain structural elements in the language rather than misunderstanding of a lexical definition; second, errors that might be describable in terms of differences in rule formulation between student and native grammar will not necessarily be explained in this way. The following section further illustrates this.

2.2 Grammatical errors that wouldn't ordinarily occur may also be induced through misunderstanding of meaning or usage, or, occasionally, through faulty explanation. Students in one advanced class were asked if they knew the meaning of *any*, and, when all said yes, to give some examples, with the following result: "In this class there are any students who speak German (=not any, no students)", "In a private garden anyone can enter (=no one)" and "Anybody has to work." Apparently these students were once told something to the effect that *any* is used in negatives, or has negative connotations, and they interpreted that to mean that it was itself a negative word, like *never* or *nothing*. Thus, in the first two sentences above, the students left the negative marker off the VP as redundant, producing, in the first case, a deviant sentence, and in the second a polarity switch. The third sentence seems to involve a confusion between *any* and *every* (cf. " $\left\{\begin{matrix}\text{Everybody}\\ \text{Anybody}\end{matrix}\right\}$ may leave now."). That these are situation-induced errors is suggested by the fact that students, even at advanced levels, tend not to use *any* and other words involving feature shifts either correctly or incorrectly, though they would probably understand them if encountered in someone else's speech. In fact, "I don't have some NP" is a common type of student error; on the other hand, I found no other cases of misuse of negative polarity words, probably because such words are among the last things learned, and hence almost never used, even at advanced levels.

One teacher defined *as if* as more or less synonymous with *like* and then asked students to transform sentences with *like* into sentences with as if (e.g., "He climbs like a monkey" into "He climbs as if he were a monkey."). But one student, given the sentence "She cries like a baby" responded with "She cries as if the baby cries" which would be fine if like were really just a synonym for as if. In fact, however, there is a structural change involved in this exercise as well. The sentence the student was asked to transform is ambiguous; it could be paraphrased in either of the following ways:

> Her crying is like that of a baby. (she cries like a baby cries)
>
> In that she cries, she is just like a baby. (she cries like she is a baby)

It is only in the second of these two readings that as if can be used. The student has inserted as if in the first sentence with incorrect results. The average native speaker of English would not normally be conscious of the above ambiguity until an incorrect sentence appeared such as the one this student produced. It is crucial, however, that the teacher be aware of such potential ambiguities and avoid inadequate explanations like the one given in this case. A thorough understanding of the syntax of sentences with like in them is the only way to make clear to the student when as if can be substituted for like and when it cannot.

A final danger lies in overreliance on grammatical terminology without sufficient attention to function in the sentence. Thus, after several examples involving absolutely impossible, students asked to use the adjective form of absolutely produced: "It's absolute impossible to do." These students had fairly thorough familiarity with grammatical terminology and usage, so it may have been sheer laziness that produced this sentence. In any case, such an error is simple enough to correct, especially where the students are acquainted with the terms. But it is not the sort of mistake they usually made spontaneously and can thus be attributed to careless reliance on terminology at the expense of the usage it represents, rather than to real confusion over form and function. As such, it is probably not worth spending much time on, and should be recognized by teachers for what it is.

2.3 It is easy to fall into the pattern of a structural drill and forget that the sentences being produced have semantic content; in so reducing textbook exercises to rote mechanical repetition, students produce some bizarre semantic violations that would be unlikely to appear in real speech. An intermediate level drill gave students the choice of several phrases as possible responses to a yes-no question: No, but I $\left\{\begin{array}{l}\text{hope to}\\\text{ought to}\\\text{must}\\\text{expect to}\\\text{have to}\\\text{etc.}\end{array}\right\}$

Thus, "Do you want to study? No, but I have to." Some of the students responded with no apparent regard for any semantic relationship between the question and their choice of response: "Are you going to the movies? No, but I hope to"; "Do you want to study tonight? No, but I hope to"; "Does she understand French? No, but she must."[10] It may have been mere lack of attention that caused such errors (only the students can know), but even those students who were listening in seemed to notice nothing peculiar in them. From the responses students were making, it appeared that they didn't really understand the exercise or the meaning of the tag they were being asked to put on their answers. Again, this sort of thing would not occur in normal speech, but probably only because they have not in fact learned the construction and hence wouldn't use it spontaneously. Such exercises need to be carefully monitored to prevent students from making a mistake set up for them by a possible wrong choice, and above all to avoid reinforcing misconceptions. Where this is not done, students are likely either to not understand or to mislearn. The first would render the drill meaningless; the second would make it detrimental.

Another sort of exercise has the opposite effect of rigidly dictating the students' responses where there are other valid possibilities. Thus, a not-insertion drill was preceded by the following instructions:

> If the negative can be contracted with an auxiliary, do so. Otherwise, put not before the infinitive.

as in: "John can write → John can't write"; "John prefers to write → John prefers not to write." In the first example above, it is the modal can which is being negated, not the verb write, whereas in the second case negation is associated with write rather than the preceding prefer. As if it weren't enough that students had two different structures to work with in the same drill, the majority of sentences of the second type involved matrix verbs of the type expect, where there are two possibilities for negative placement, with a corresponding difference in scope and meaning. Moreover, the "correct" response was often the less common of the two possibilities. Students were expected to say

John $\left\{\begin{array}{l}\text{expects}\\ \text{wants}\\ \text{etc.}\end{array}\right\}$ not to go.

with the predictable result that they were torn between what they were instructed to say and what they had so often heard and tried to imitate: "John doesn't expect to go." In addition, they had to cope with input sentences like "He will continue to answer" where both contraction-after-modal or not-before-infinitive are possible, and "He needs to study" with its variety of possibilities for negation: "He $\left\{\begin{array}{l}\text{doesn't need to}\\ \text{needs not to}\\ \text{need not (needn't)}\end{array}\right\}$ study" the latter in its uncontracted form being the one taught, for no discernible reason and with no explanation.[11]

A third potential source of trouble in exercises can be found in those drills which involve the joining of two simple sentences in which grammatical transformations have applied which would not apply in the complex sentence. Thus, for example, students were asked to embed one of two sentences in the other by using one of the conjunctions unless, because, if, although, whenever, whether or not, as in the following:

The air conditioner isn't working. The students enjoy the class.

$\Rightarrow$ The students enjoy the class $\left\{\begin{array}{l}\text{unless}\\ \text{although}\\ \text{etc.}\end{array}\right\}$ the air conditioner isn't working.

Taking an imperative sentence as part of his input, one student produced the following: "I can't buy any new shoes unless lend me some money" (from "Lend me some money. I can't buy any new shoes."). That is, he joined the surface strings, as they were given to him, and of course imperatives can't occur in the resulting environment.

What is interesting is not so much this mistake (which is not surprising and is easy to correct) as the fact that in other cases where students were thus given an opportunity to make a mistake from the input to their drill, they did not. One pair of sentences given in the same exercise as that just discussed was this: "Paul is too busy. He takes a nap every afternoon." Joining these with unless, one would expect "*He takes a nap every afternoon unless Paul is too busy" which is a violation of the rules of pronoun reference. In fact, however, this did not occur. Students changed the sentence to avoid this deviance: "He takes a nap every afternoon unless he is too busy" or "Paul takes a nap every afternoon unless he is too busy" and in other cases, students were able to form various complex sentences involving pronominalization always keeping the pronoun-antecedent relations straight. A few wild sentences involving co-reference did occur: "Girls are in contact with many young men; at first they become friends, and if a girl falls in love with one of them, they will agree to be married." I am not sure this is all that bad; there are some reference switches, but this is the sort of sentence which occurs in native speech all the time, and as such differs from what one generally means in speaking of "student errors". It is interesting to note, on the other hand, that the sorts of mind-boggling sentences linguists play with never seem to occur in the speech of learners of English, and it is worth asking why not.[12]

3. Unexpected Errors

3.1 Common errors. There were other types of errors that students insisted on making, despite strong inducements to avoid them. For example, students at all levels often substituted his for her and her for their, both in free speech and in reading, when the correct form was right in front of them. They had particular trouble with the final s of the plural (noun) and third person singular (verb) and often left it out, even when reading, as in "How's the work these day, Jack?" (from a reading exercise).[13] An example of just how difficult the -s ending is can be seen in the following: in a simple substitution drill where students were to change the subject of a sentence as indicated by the teacher and change the form of the verb where necessary, the following exchange took place:

Student$_1$: He is often here.
Teacher: The students
Student$_2$: The student is often here.
Teacher: Studentsssss
S$_2$: They are often here.
T: Say studentssssss
S$_2$: The student are often here.

This student clearly knew, at least by the end of the exchange, that a plural was involved, and knew how to form the plural for the verb, but try as he would the teacher could not get her to pronounce the plural morpheme -s on the noun.

Tense was another frequent problem, and particularly certain sequence of tense rules. Thus, there were mistakes such as "[In the old days, parents arranged weddings and] the boy didn't know the girl he is going to marry" and "After a long travel, I returned to my house and see a lot of dust." The following sentence "He acts as if he knows Habib Bourguiba" was given in response to the exercise changing like to as if (contrary to fact). The students had been told repeatedly that as if requires a past verb for a counterfactual meaning, and in this case the input sentence given by the teacher already had the correct verb form in it, ("He acts like he knew H.B."), so that there was no change or choice that the student needed to make to produce the expected "He acts as if he knew Habib Bourguiba". Yet he changed the tense, and thus the meaning of the sentence. In cases like this, apparently the student's difficulty with the form overcame all the positive clues he was given.

In general, students had far less trouble with the past tense in simple sentences, although errors such as "I forget to say to you that..." and "Do they have lunch yesterday" did occur on a random basis, even among advanced students.

3.2 Atypical errors. Not all errors made in these circumstances are typical in other contexts, however. A common instance of a situation where all the right information is available is the pattern drill, where students have just one transformation to perform, leaving the rest of the sentence intact. One of the purposes of this sort of drill is precisely to militate against extraneous, irrelevant errors, and as has been discussed above, students tend to take this sort of instruction quite literally, and where an exercise is not carefully written, may in fact produce a deviant sentence by choosing the wrong lexical item to insert or by failing to make some secondary change which goes along with the transformation being drilled. But where there is truly only one change to be made, one expects the student to leave the rest of the sentence alone (and herein lies the ease in ignoring its meaning altogether). It is therefore surprising to find a student reordering other elements in a sentence in the course of a drill, even where word order may be troublesome elsewhere in his speech. In a passive formation drill, a student given "We always keep our knives in this drawer" produced the following: "Our knives are kept in this drawer always." While some students have difficulty with the order of constituents in English, adverb placement, at least with simple tenses, is not generally a troublemaker. The fact that the simple present is changed to an Auxiliary+ participle construction may be the source of trouble, but what one would expect, on the basis of similar spontaneously made errors, would be something like "Our knives are kept always in this drawer." I know of no precedent or explanation for reordering the adverb with respect to the prepositional phrase in the course of the drill.

In a dialogue that was being learned by repetition and memorization, the following sentences were produced: "Ours is much newer model," and "His must be a one of the newest models." In a similar case, students were forming sentences on the pattern "Verb+ing is fun," "X-ing is hard work"; one student said "Walking is a good exercise." The first of these sentences represents one of the most common mistakes made by Tunisian students--omission of the article. In the next two cases, however, an article has been inserted, something far rarer in free conversation. These may be cases of overcorrection; a grammatical explanation could no doubt also be found. It is peculiar, however, that this occurs here, where the model for the correct sentence without the article is available to the student and is the more natural one for him in any case.

The above error types represent an entirely different situation from those described in earlier sections. The sentences discussed in 3.1 are examples of patterns that also occur frequently in spontaneous speech, and which are so strongly a part of the student's grammar that they are preferred even where the correct form is obvious.[14] The examples in this section, on the other hand, represent a peculiar subset of induced errors. Like induced errors, the cases describe above do not occur in normal speech, but here the situation in the classroom is set up not for them to be made, but for them NOT to be made. Yet they occur despite everything. Such cases are totally baffling within the framework of induced errors.

4. The Problem of Meaning

It is often unclear what a student means to say when he produces an ungrammatical sentence, particularly when it has no context, as in answer to a teacher's request to "give me a sentence using X." One student, when asked to give a sentence using as if (which had just been defined as a synonym of like with the second clause in the past for negative implication), said "I am glad as if I had slept." It is not at all clear what is wrong with this sentence because the meaning that the student had in mind is obscure (does he mean: glad because I slept, glad that I slept, would be glad if I'd slept, would like to have slept, didn't sleep but feel as if I had?). Since the teacher didn't ask the student to explain, he could not do more than just say "you can't say that in English."

In a lesson on gerunds, students were questioned in such a way that they produced sentences with gerundive subjects:

Teacher: What do you do on the weekend?
Student$_1$: I go swimming.
T: Is it fun?
S$_1$: Yes.
T: What's fun?
S$_2$: Swimming is fun.

One such sequence went as follows:

T: What does your father do?

S_1: He's director of a bank.

T: Is it hard work?

S_1: Yes.

T: What's hard work?

S_2: Hard work is the director of a bank.

This is an example of the error type discussed in section 3.2 above, where the student ignores a drill situation which sets him up for the correct response, and produces a deviant sentence as a result. The clue to what might be happening in these cases must be sought in the student's intended meaning. One could hypothesize several different possibilities for this sentence, taken in isolation. Notice, for example, that the above response is identical in form to a different type of WH-question-answer pair:

Q: What is an apple?

A: An apple is a fruit.

where the questioned noun phrase subject is defined by qualification in the predicate. That is, in one case, the predicate provides the new information, and in the other, the subject does. The question here is whether the student really has confused the two: did he misunderstand the question or simply not know how to form the answer? It seems unlikely in a case where several correct sentences on the same pattern had already been elicited that he would suddenly do a semantic reverse like this. In fact, we have already seen that evidence seems to indicate that the opposite is more often the case. That is, if in the middle of a series of sentence pairs of the pattern "What is fun? Swimming is fun," the teacher said "What is an apple?", a student with poor command of wh-questions might respond, in the same pattern, with "Fruit is an apple." The pattern is not broken, even when it should be. It is more likely in this instance that the student understood the question and had the correct answer in mind. It was only coincidental that, in garbling the grammar of his answer, he stumbled upon another English syntactic pattern. The question then remains, why did he make a mistake, and what should the teacher do about it? Is it enough simply to give the right answer, with no explanation of why the sentence was wrong? Or should the teacher take the time to find out exactly what the student meant, and show him why he didn't succeed in expressing it? If his problem is something as simple as not knowing what to do with the object of the gerund, merely learning the correct sentence might be enough. But the teacher must know his student well to be sure the problem isn't deeper, especially where the student's sentence falls so close to another correct English pattern and where the drill situation would lead one to expect the right answer. Even if he didn't have the definition-type sentence in mind, he might still notice the resemblance later and get confused. Alternatively, the student may, as has been suggested by Corder (1967), simply be testing a hypothesis about the new language by the sentence he has produced and will accept the correction. The only way to tell in such cases is by studying the persistence of errors over time. Only the teacher, who sees his students regularly over a period of time, is in a position to know what is necessary for this part of error analysis.

It may not even be clear in every case whether a sentence a student produces is deviant or not. For example, in a class asked to make up sentences using should, one student said, "They should need another book." If should is taken to mean expectation on the part of the speaker, this sentence is fine ("They should need another book by next week, because they're on the last chapter of this one already"). But it is unlikely that this is what the student meant, since this class was just learning modals for the first time, and should, even in its more common sense of ought to was not an active part of their vocabulary. On the other hand, co-occurrence restrictions such as these are seldom violated in spontaneous speech, and it is unlikely that a sentence such as this would have been produced if the student had not been put on the spot. In a situation like this, the teacher is faced not only with a decision about the student's meaning but with the question of whether to point out the possible correct reading and perhaps confuse a student who hadn't even considered the alternative.

The time spent in trying to figure out what a student meant may not always be justifiable. But it is important for the teacher to be aware of all the possibilities in order to make this decision. Correcting the wrong thing, or the right thing for the wrong reason, or not correcting enough, can easily make matters worse.

5. Summary and Conclusions

5.1 In the weak version of contrastive analysis, as in all branches of error analysis, the source of the error is an important issue (unlike the strong version of contrastive analysis, where the source is assumed). A great deal of evidence has come to light indicating that contrastive analysis is inadequate to explain the source of certain types of error. My research in Tunisia has reaffirmed this. But error analysis, as it is usually defined (that is, as an alternative explanation to contrastive analysis) also begs the question. Too many variables are involved to say that the idiosyncracies of learner language are explainable solely in terms of either interference or problems internal to the target language. In this paper, I have concentrated on errors which are related to the classroom situation itself. These fall into several categories. The first, and most obvious, are the errors which students wouldn't make in free speech, but which are elicited by the teacher's questions or by drills. These may be due to incomplete acquisition of the lexical item or grammatical structure involved, to analogy suggested by the order of presentation, or to a number of other possible causes, as yet unexplored.

I also touched briefly on some errors which one might expect to find within this framework, but which students manage to avoid. The relationship of potential errors which do not actually occur to putative linguistic universals such as coreference restrictions, semantic restrictions, meaning conflicts, etc. is worth exploring. To this end, a thorough study of what errors are not made will be necessary. In addition there are errors that recur despite a teaching situation designed to avoid them. These present different problems of source identification, and are really a separate category. These errors may provide evidence for the hypothesis-strategy approach to language learning suggested by Corder (1971). If such is the case, it may turn out not to be worthwhile to force extensive drill of these issues when the learner isn't ready to absorb them.

5.2 The discussions above are intended to be highly tentative, as evidenced by the fact that no attempt is made to provide any solutions. It is first necessary to point out the possible sources of error, in order to make teachers aware of the problems. Doubtless induced errors will have to be treated differently in the classroom from those which actually reflect the student's developing competence, but it is not clear just how this is to be done.

Several issues are involved here. In preparing materials and writing drills, it is important to keep in mind that surface relationships could lead to misunderstanding on a deeper level, and that any inexplicit explanation may be open to misinterpretation by the student. The teacher must be on guard to monitor drills and explanations to be certain that they do not lead to false generalizations. In this way, many induced errors could be avoided or caught before they cause any real trouble. There is also the question of what to do about those errors which will crop up and persist despite all attempts to control them. It is here that the issue of ascertaining the student's meaning plays an important role. The clearer a teacher's understanding of the sources of student errors, the better he will be able to judge which ones are most worth concentrating on. Finally, an understanding of the effect on the student of miscorrection through misinterpretation of his intent is vital to adequate language teaching. No teacher would disagree that correcting the wrong thing could be disastrous; however, it is not always obvious what is the right and what is the wrong thing. Only after this is determined, and only over an extended period of time, can the nature and extent of damage that such cases might cause be measured.

For the linguist studying second-language acquisition, the distinction between teacher- or situation-induced errors and spontaneous student errors is also valuable. It is crucial to bear this distinction in mind when attempting to account for student language, for to ignore it is to risk faulty analysis. It could cloud the issues to attribute to a developing grammar an error that the student wouldn't ordinarily make. It is of course quite possible to describe many, and perhaps all, of the errors discussed above solely in terms of rule deviance. For example, one could describe the confusion between point out and notice described in section 2.1 as simply a patient-agent confusion in the student's grammar, or the inability to distinguish between should and must by formally giving them a slightly different meaning from that of standard English. But such statements would be misleading if the description of student language is to have any reality in terms of the student's acquisition of the language, for these errors do not in fact represent any internalized part of his grammar, and do not therefore occur systematically in the same way that certain errors of tense usage or word order might occur. "Whenever a student of a second language creates and utterance in the second language, he reveals something about his competence (or his lack of competence)" (Wardhaugh, 1967). The errors discussed above suggest something about what the student does not know rather than about what he does know. We must not make the mistake of attributing them to a system where none exists. An approach to student language description which does not distinguish between spontaneous, systematic errors and those forced by the situation will do nothing to explain where the student is in the second language acquisition process, how he got there, or what the process is.

FOOTNOTES

This work was supported in part by Contract OEC-0-70-4986 from the U.S. Office of Education to the Language Research Foundation. I am endebted to many people who helped me in my research both here and in Tunisia. I would like especially to express my thanks to Mohamed Maamouri of the Institut Bourguiba des Langues Vivantes, Tunis, who arranged for me to attend English classes there, to DeWayne Coambs, Leslie and Lewis Tremaine, Mike Vale, and other Peace Corps staff members for access to their classrooms, and for helpful discussions of their own experiences at the beginning stages of my research, and to Bruce Fraser, Walt Olson, Patty Regan and John Schumann for their valuable comments on earlier versions of this paper.

1. Specifically, this version of contrastive analysis "demands of linguists that they have available a set of linguistic universals formulated within a comprehensive linguistic theory which deals adequately with syntax, semantics, and phonology. Furthermore, it requires that they have a theory of contrastive linguistics into which they can plug complete linguistic descriptions of the two languages being contrasted so as to produce the correct set of contrasts between the two languages." (Wardhaugh (1970) p. 125.) There is also the problem of needing different materials for every native language group, since the predicted difficulties would presumably be different for each. It is trying to work without such prerequisites that leads to the oft-levelled criticism of superficiality.

2. This is in line with Bowen's (1967) claim that contrastive analysis is not a method, but a technique for organizing facts. I shall return to this issue in my discussion of error analysis as well.

3. By performance errors I mean errors of the sort which even native speakers make, due to memory lapses, inattention, etc. Thus, the distinction between competence and performance in the sense of Chomsky (1965) must be kept in mind in studying second language acquisition as well as in linguistic theory and first language acquisition. We shall not be concerned with problems of performance in this paper.

4. Of course, it has been claimed that contrastive analysis would be redeemed if only a deeper level analysis is used for the contrast, but that only leaves the analyst vulnerable to the criticism that the prediction depends wholly on the analysis. The more abstract the analysis, the less reliable it gets, given the present state of linguistic knowledge. Such an analysis must await the further development of linguistic theory.

5. In fact, different people tend to mean quite different things when they speak of error analysis. In contrast to Buteau and Duškova, who seem to consider it as an alternative or supplement to contrastive analysis, Corder (1971) suggests that errors be considered not as rule violations but as rule differences between divergent dialects of a given language. As such, they serve as clues to the system the learner has constructed for the language he is studying, and the ways in which it differs from the target dialect.

6. In order to do this, the teacher would have to be able on many occasions to read his students' minds, an unrealistic expectation under normal circumstances. I discuss this problem somewhat further, though by no means conclusively, in section 4. Notice that the three steps of error analysis described by Corder (1971) -- identification, description, and explanation -- make no reference to use in actual teaching. Such application would have to follow accomplishment of the other three steps.

7. Error types and sources have been grouped in other ways by various people. For example, Richards (1971) quotes Selinker as hypothesizing five central processes in second language learning: reorganization of linguistic materials, language transfer (interference), strategies of learning, strategies of communication, and transfer-of-training. Elsewhere references have been made to generalization, interference of form and function, overcompensation, and others, which seem to be more specific examples of the first of Selinker's processes. For my purposes I have collapsed his first four processes under the general heading "incomplete target language acquisition"; transfer-of-training is the source of the induced errors I discuss herein, and I have added the performance factor as my third category. Thus there is no real conflict between my categorizations and Selinker's. All such categorizations of learning processes should be held open as possibilities until the ways in which they fit together are better understood; as they have been little studied so far, many may overlap in ways which we do not now forsee.

8. Thus, this error falls into the category of those discussed in section 3 -- those which persist despite correction. I include it in this section because it is such a clear example of a lexical misuse induced by analogy to its definition. Not all the examples of section 3 are so nicely describable in terms of this section as well.

9. The confusion lies in agent-patient relationships. For notice, the agent is also the patient, whereas this is not the case for point out. There are other instances of words which may be a similar source of confusion--especially since such relationships are by no means universal, but vary from language to language. The teacher of any foreign language should be at all times acutely aware of precisely what sorts of features are involved. This will prove crucial to an understanding of student errors -- which sorts can be induced and which can't and how to treat different kinds.

10. This latter example is not necessarily an example of a semantic violation; in fact it may reflect a common spontaneous error type among students at this level: the substitution of must for should. Note that a response "No, but she should" would have been perfectly acceptable. This may well be what the student was saying. The same may be true of other examples I have used. This points up once again the extreme tentativeness of any claim about what the student actually meant to say. See also section 4 on this subject.

11. This is a classic example of the deep/surface structure conflict: two sentences which have the same structure on the surface are quite different at a deeper, more abstract level.

An awareness of current research in transformational linguistics can help the teacher find and avoid such areas of difficulty. 12. Clues to the question of which possible errors students avoid despite inducements to make them may lie in the notion of linguistic universals. The notion of command relationships and the constraints on pronominalization associated with them have been proposed as universal to all languages. See Langacker (1969) for a discussion of the notion of command. Bever (1970) has formulated this as a perceptual constraint to the effect that one element cannot stand for another unless a connection has already been established or it is apparent (as through a marker of subordination) that such a connection is about to be established. In either case, it would not be surprising for a student to avoid violating a universal constraint. Imperative formation, on the other hand, differs from language to language (although some form of imperative is universal). Thus, Arabic does not distinguish imperative sentences from declaratives by subject deletion as does English. It would be interesting to examine whether students whose native language does form imperatives by subject deletion would make the same mistake. This is an area where theoretical linguistics and teaching can perhaps be mutually useful, the research on universals (syntactic or cognitive) providing a possible explanation for the occurrence and non-occurrence of certain errors, and the classroom providing a place to test hypotheses about universals.

13. This is perhaps the most widespread phenomenon I have observed, occurring in the speech of students of many different language backgrounds at all levels.

14. That is, unlike the bulk of examples in this paper, there appears to be nothing in the classroom situation which would explain these errors. They seem to constitute a thoroughly internalized part of the student's grammar. It has been suggested (Corder, 1971) that such cases not be considered errors at all, but rather manifestations of rule differences between a student's idiosyncratic dialect and the native dialect he is learning. Insofar as these cases do not fit into the pattern of errors discussed elsewhere in this paper, they may constitute evidence supporting this position.

References

Banathy, Bela H., and Paul H. Madarasz (1969) "Contrastive Analysis and Error Analysis" Journal of English as a Second Language, IV:2, pp. 77-92.

Bever, Thomas G. (1970) "The Cognitive Basis for Linguistic Structures" in Hayes, John R. (ed.) Cognition and the Development of Language, New York & London: John Wiley & Sons, Inc.

Bowen, J. Donald (1967) "Contrastive Analysis and the Language Classroom" in Robinett, Betty Wallace (ed.) On Teaching English to Speakers of Other Languages, Series III, Papers read at the TESOL Conference, New York, March 17-19, 1966. Washington, D.C.: TESOL.

Burt, M.K. and Carol Kiparsky (forthcoming) The Gooficon, Rowley, Mass.: Newbury House.

Buteau, Magdelhayne F. (1970) "Students' Errors and the Learning of French as a Second Language: A Pilot Study" International Review of Applied Linguistics and Language Teaching (IRAL) VIII:2, pp. 133-145.

Chomsky, Noam (1965) Aspects of the Theory of Syntax, Cambridge, Mass.: MIT Press.

____________ (1966) "Linguistic Theory" in Robert G. Mead, Jr., (ed.) Language Teaching: Broader Contexts, Northeast Conference Reports, 1966.

Corder, S. P. (1967) "The Significance of Learners' Errors", IRAL V:4, pp. 161-169.

Corder, S. P. (1971) "Idiosyncratic Dialects and Error Analysis" IRAL IX:2, pp. 147-159.

Duškova, Libuše (1969) "On Sources of Errors in Foreign Language Learning" IRAL VII:1, pp. 11-36.

Jacobs, Roderick (1969) "Linguistic Universals and Their Relevance to TESOL" TESOL Quarterly III:2, pp. 117-122.

James, Carl (1969) "Deeper Contrastive Study" IRAL VII:2.

Langacker, Ronald W. (1969) "Pronominalization and the Chain of Command" in D. Reibel and S. Schane (eds.) Modern Studies in English, Englewood Cliffs, N.J.: Prentice Hall.

Richards, Jack C. (1970) "A Non-Contrastive Approach to Error Analysis" ERIC, ED 037 721, paper presented at TESOL Convention, San Francisco, March 1970.

________________ (1971) "Error Analysis and Second Language Strategies" Language Sciences No. 17, Bloomington, Indiana, pp. 12-22.

Rutherford, William E. (1968) "Deep Structure, Surface Structure, and Language Drill" TESOL Quarterly II:2, pp. 71-79.

Strevens, Peter (1969) "Two Ways of Looking at Error Analysis" ERIC, ED 037 714.

Wardhaugh, Ronald (1967) "Some Current Problems in Second Language Teaching" Language Learning XVII:1&2, pp. 21-26.

________________ (1970) "The Contrastive Analysis Hypotnesis" TESOL Quarterly IV:2, pp. 123-129.